

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 23, 2001, 13:22:06 ; Search time 20.3 Seconds  
(without alignments)  
907.997 Million cell updates/sec

Title: US-09-483-831-69  
Perfect score: 5019  
Sequence: 1 MARRSSFQSQIISLTFVAV.....RSPYELTLKTYLHYESEI 915

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5019	100.0	915	1	US-08-346-455B-69
2	5019	100.0	915	3	US-08-977-221-69
3	5019	100.0	915	5	PCT-US95-06613-69
4	4533	90.3	829	1	US-08-346-455B-34
5	4533	90.3	829	3	US-08-977-221-34
6	4533	90.3	829	5	PCT-US95-06613-34
7	4418	88.0	861	1	US-08-346-455B-67
8	4418	88.0	861	3	US-08-977-221-67
9	4418	88.0	861	5	PCT-US95-06613-67
10	4279	85.3	979	1	US-08-346-455B-38
11	4279	85.3	979	3	US-08-977-221-38
12	4279	85.3	979	5	PCT-US95-06613-38
13	3842	76.5	788	1	US-08-346-455B-36
14	3842	76.5	788	3	US-08-977-221-36
15	3842	76.5	788	5	PCT-US95-06613-36
16	1955.5	39.0	873	3	US-09-187-331-6
17	1955.5	39.0	925	2	US-08-392-946-1
18	1955.5	39.0	925	2	US-08-504-169-1
19	1955.5	39.0	925	5	PCT-US94-14893-1
20	360	7.2	438	3	US-09-187-331-2
21	357	7.1	355	2	US-09-014-969-19
22	192	3.8	39	1	US-08-346-455B-54
23	192	3.8	39	3	US-08-977-221-54
24	192	3.8	39	5	PCT-US95-06613-54
25	126.5	2.5	2471	4	US-09-112-450-4
26	123	2.5	4654	4	US-08-476-515A-84
27	123	2.5	4655	4	US-08-652-877-84

28	123	2.5	4655	4	US-08-652-877-88	Sequence 88, Appl
29	123	2.5	4655	4	US-08-652-877-90	Sequence 90, Appl
30	122	2.4	4655	4	US-08-652-877-86	Sequence 86, Appl
31	118	2.4	23	1	US-07-822-043-11	Sequence 11, Appl
32	115	2.3	1833	3	US-08-479-722B-2	Sequence 2, Appl
33	115	2.3	1833	5	PCT-US95-02251-18	Sequence 18, Appl
34	113	2.3	22	1	US-07-822-043-33	Sequence 33, Appl
35	113	2.3	22	1	US-08-346-455B-33	Sequence 33, Appl
36	113	2.3	22	3	US-08-977-221-33	Sequence 33, Appl
37	113	2.3	22	5	PCT-US95-06613-33	Sequence 33, Appl
38	112	2.2	1394	6	517197-30	Patent No. 517197
39	110	2.2	23	1	US-08-346-455B-11	Sequence 11, Appl
40	110	2.2	23	3	US-08-977-221-11	Sequence 11, Appl
41	110	2.2	23	5	PCT-US95-06613-11	Sequence 11, Appl
42	109.5	2.2	1810	5	PCT-US95-11684-4	Sequence 4, Appl
43	108	2.2	465	3	US-08-701-582D-13	Sequence 13, Appl
44	108	2.2	465	3	US-09-063-893A-19	Sequence 19, Appl
45	107.5	2.1	605	4	US-09-042-785A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-08-346-455B-69  
; Sequence 69, Application US/08346455B  
; Patent No. 5731167  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA; DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,455B  
; FILING DATE: 28-NOV-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 915  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: cDNA

```

; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: A2058 ATX protein
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-08-346-455B-69

Query Match      100.0%; Score 5019; DB 1; Length 915;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARRSSFQSCQIIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNIGSCK 60
Db 1 MARRSSFQSCQIIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNIGSCK 60

QY 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCKLTARGWECTKRCGEVRNEENACHC 120
Db 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCKLTARGWECTKRCGEVRNEENACHC 120

QY 121 SEDCLARGDCCTNYQVVKGESHVWDDDCCEEIKAAECPAGFVRPPLIIFSVDGFRASYMK 180
Db 121 SEDCLARGDCCTNYQVVKGESHVWDDDCCEEIKAAECPAGFVRPPLIIFSVDGFRASYMK 180

QY 181 KGSKVMPNIEKLRSCTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240
Db 181 KGSKVMPNIEKLRSCTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240

QY 241 TFHLRGREKFNHRWGGQPLWTATKQGVKAGTFEWSVVIPHERRILTLRWLTLPDHER 300
Db 241 TFHLRGREKFNHRWGGQPLWTATKQGVKAGTFEWSVVIPHERRILTLRWLTLPDHER 300

QY 301 PSVYAFYSEQPDFSGHKYGPFGPESSYSGSPFPAKRPRKRVAPKRRQERPVPAPKKRRR 360
Db 301 PSVYAFYSEQPDFSGHKYGPFGPESSYSGSPFPAKRPRKRVAPKRRQERPVPAPKKRRR 360

QY 361 KIHRLMDHYAAETRODKMTNPLREIDKIVGQMDGLKQLKRRRCNVNIFVGDHGMEDVTCD 420
Db 361 KIHRLMDHYAAETRODKMTNPLREIDKIVGQMDGLKQLKRRRCNVNIFVGDHGMEDVTCD 420

QY 421 RTEFLSNLYTNVDDITLVPGLTGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQH 480
Db 421 RTEFLSNLYTNVDDITLVPGLTGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQH 480

QY 481 LPKRLHYANNRRIEDIHLLVERHWVARKPLDVYKKPSGKCFQGDHGFNDKNVSMQTVE 540
Db 481 LPKRLHYANNRRIEDIHLLVERHWVARKPLDVYKKPSGKCFQGDHGFNDKNVSMQTVE 540

QY 541 VGYGPTFKYKTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRTNTFRPTMPEEV 600
Db 541 VGYGPTFKYKTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRTNTFRPTMPEEV 600

QY 601 TRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDLNLKRLHTKGTSTERRHLLYGRPAVLRYT 660
Db 601 TRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDLNLKRLHTKGTSTERRHLLYGRPAVLRYT 660

QY 661 RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDRVSPSFSQNCL 720
Db 661 RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDRVSPSFSQNCL 720

QY 721 AYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPFAKRVWNYFQRLVKKYASER 780
Db 721 AYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPFAKRVWNYFQRLVKKYASER 780

QY 781 NGVNVISGPIFDYDYDGLHDTEDKIQYVEGSSIPVPTHYYSIIITSCLDFTQPADKCDGP 840
Db 781 NGVNVISGPIFDYDYDGLHDTEDKIQYVEGSSIPVPTHYYSIIITSCLDFTQPADKCDGP 840

QY 841 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSSRYPE 900
Db 841 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSSRYPE 900
```

```

QY 901 ILTLKTYLHYESEI 915
Db 901 ILTLKTYLHYESEI 915

RESULT 2
US-08-977-221-69
; Sequence 69, Application US/08977221
; Patent No. 6084069
; GENERAL INFORMATION:
; APPLICANT: UNITED STATES OF AMERICA; DEPT.
; APPLICANT: OF HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: MOTILITY STIMULATING
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,221
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,455
; FILING DATE: 28-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,182
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/822,043
; FILING DATE: 17-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4149US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: Unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: No
; FEATURE:
; NAME/KEY: A2058 ATX protein
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-08-977-221-69
```

```

Query Match      100.0%; Score 5019; DB 3; Length 915;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARRSSFQSCQIIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNIGSCK 60
Db 1 MARRSSFQSCQIIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNIGSCK 60

QY 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCKLTARGWECTKRCGEVRNEENACHC 120
```

Db 61 GRCFELQAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCEVRNEENACHC 120  
QY 121 SEDCLARGDCCTNYQVCKGESHVWDDCEIEKAAECAPAGVVRPPLIIFSVDFGRASYMK 180  
Db 121 SEDCLARGDCCTNYQVCKGESHVWDDCEIEKAAECAPAGVVRPPLIIFSVDFGRASYMK 180  
QY 181 KGSVMPNIEKLRSCGTHSPYMRPVYPTKTFPNTLYLATGLYPESHGIVGNSMYDPVFDA 240  
Db 181 KGSVMPNIEKLRSCGTHSPYMRPVYPTKTFPNTLYLATGLYPESHGIVGNSMYDPVFDA 240  
QY 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPIHERRILTLRWLTLPDHER 300  
Db 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPIHERRILTLRWLTLPDHER 300  
QY 301 PSVYAFYSEQPDFSGHKYGFPGPEESSYGSPTPAKRPKRKVAPKRRQRPVAPPKRRR 360  
Db 301 PSVYAFYSEQPDFSGHKYGFPGPEESSYGSPTPAKRPKRKVAPKRRQRPVAPPKRRR 360  
QY 361 KIHRMDHYAAETRODKMTNPLREIDKIVGQMLDGLKQLRRCVNVIFVGDHGMEDVTCD 420  
Db 361 KIHRMDHYAAETRODKMTNPLREIDKIVGQMLDGLKQLRRCVNVIFVGDHGMEDVTCD 420  
QY 421 RTEFLSNLYTNVDDITLVPGTGLGRIRSKFSNNAKYDPAKIAIANLTCKKPDQHFKPYLKQH 480  
Db 421 RTEFLSNLYTNVDDITLVPGTGLGRIRSKFSNNAKYDPAKIAIANLTCKKPDQHFKPYLKQH 480  
QY 481 LPKRLHYANNRRIEDIHLVRRWHVARKPLDVYKPSGKCFQGDHGFNDKVNMSQTVF 540  
Db 481 LPKRLHYANNRRIEDIHLVRRWHVARKPLDVYKPSGKCFQGDHGFNDKVNMSQTVF 540  
QY 541 VGYGPTFKYKTYPPFENIELYNVCDLLGLKPAPNNGTHGSLNHLRTNTFRPTMPEEV 600  
Db 541 VGYGPTFKYKTYPPFENIELYNVCDLLGLKPAPNNGTHGSLNHLRTNTFRPTMPEEV 600  
QY 601 TRPNYPGIMYLOSDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEEHLLYGRPAVLRYT 660  
Db 601 TRPNYPGIMYLOSDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEEHLLYGRPAVLRYT 660  
QY 661 RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPFSQNC 720  
Db 661 RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPFSQNC 720  
QY 721 AYKNDKQMSYGLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRWVWNYFORVLVKYASER 780  
Db 721 AYKNDKQMSYGLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRWVWNYFORVLVKYASER 780  
QY 781 NGVNVISGPIFDYDGLDHTEDKIKQYVEGSSIPVTHYYSIITSCLDFTQPADKCDGP 840  
Db 781 NGVNVISGPIFDYDGLDHTEDKIKQYVEGSSIPVTHYYSIITSCLDFTQPADKCDGP 840  
QY 841 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSTRYPE 900  
Db 841 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSTRYPE 900  
QY 901 ILTLKTYLHTYESEI 915  
Db 901 ILTLKTYLHTYESEI 915

RESULT 3  
PCT-US95-06613-69  
; Sequence 69, Application PC/TUS9506613  
; GENERAL INFORMATION:  
; APPLICANT: SPRACKE, MARY; LIOTTA, LANCE;  
; APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,  
; APPLICANT: HENRY; MURATA, JUN  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/346,455  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US2  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 915  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: No  
FEATURE:  
NAME/KEY: A2058 ATX protein  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PCT-US95-06613-69

Query Match 100.0%; Score 5019; DB 5; Length 915;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MARRSSFQSQIISLFTFAVGVSICLGTAAHRIKRAEGWEEGPTVLSDPWTNIGSCK 60  
Db 1 MARRSSFQSQIISLFTFAVGVSICLGTAAHRIKRAEGWEEGPTVLSDPWTNIGSCK 60  
QY 61 GRCFELQAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCEVRNEENACHC 120  
Db 61 GRCFELQAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCEVRNEENACHC 120  
QY 121 SEDCLARGDCCTNYQVCKGESHVWDDCEIEKAAECAPAGVVRPPLIIFSVDFGRASYMK 180  
Db 121 SEDCLARGDCCTNYQVCKGESHVWDDCEIEKAAECAPAGVVRPPLIIFSVDFGRASYMK 180  
QY 181 KGSVMPNIEKLRSCGTHSPYMRPVYPTKTFPNTLYLATGLYPESHGIVGNSMYDPVFDA 240  
Db 181 KGSVMPNIEKLRSCGTHSPYMRPVYPTKTFPNTLYLATGLYPESHGIVGNSMYDPVFDA 240  
QY 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPIHERRILTLRWLTLPDHER 300  
Db 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPIHERRILTLRWLTLPDHER 300  
QY 301 PSVYAFYSEQPDFSGHKYGFPGPEESSYGSPTPAKRPKRKVAPKRRQRPVAPPKRRR 360  
Db 301 PSVYAFYSEQPDFSGHKYGFPGPEESSYGSPTPAKRPKRKVAPKRRQRPVAPPKRRR 360



QY 361 KIHMDHYAAETRDQKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTC 420  
Db 361 KIHMDHYAAETRDQKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTC 420  
QY 421 RTEFLSNLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFYPYLKQH 480  
Db 421 RTEFLSNLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFYPYLKQH 480  
QY 481 LPKRLHYANNRRIEDIHLVVERRWHVARKPLDVYKKPSGKCFQGDHGFNKNVSMQTVF 540  
Db 481 LPKRLHYANNRRIEDIHLVVERRWHVARKPLDVYKKPSGKCFQGDHGFNKNVSMQTVF 540  
QY 541 VGYGPTFKYTKVPPFENIELYNVCMCDLLGLKPAPNNGTHGSLNHLRTNFRPTMPEEV 600  
Db 541 VGYGPTFKYTKVPPFENIELYNVCMCDLLGLKPAPNNGTHGSLNHLRTNFRPTMPEEV 600  
QY 601 TRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRT 660  
Db 601 TRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRT 660  
QY 661 RYDILYHTDFESGYSEIFLMLLTWTSYTVSKQAEVSSVPDHLTSCVRPDVRSVPSFSONCL 720  
Db 661 RYDILYHTDFESGYSEIFLMLLTWTSYTVSKQAEVSSVPDHLTSCVRPDVRSVPSFSONCL 720  
QY 721 AYKNDKQMSYGFLEPPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRLVKKYASER 780  
Db 721 AYKNDKQMSYGFLEPPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRLVKKYASER 780  
QY 781 NGVNVISGPIFDYDGLDHTEDKIKQYVEGSSIPVPTHYYSIITSLDFTQPADKCDGP 840  
Db 781 NGVNVISGPIFDYDGLDHTEDKIKQYVEGSSIPVPTHYYSIITSLDFTQPADKCDGP 840  
QY 841 LSVSSFILPHRPDNEESCNSSEDESKWEELMKMHTARVRDIEHLTSLDFFRKTSRSTYPE 900  
Db 841 LSVSSFILPHRPDNEESCNSSEDESKWEELMKMHTARVRDIEHLTSLDFFRKTSRSTYPE 900  
QY 901 ILTLKTYLHTYESEI 915  
Db 901 ILTLKTYLHTYESEI 915

RESULT 4  
US-08-346-455B-34  
; Sequence 34, Application US/08346455B  
; Patent No. 5731167  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA; DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,455B  
; FILING DATE: 28-NOV-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 829  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: No  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE: Melanoma  
; CELL LINE: A2058  
; ORGANELLE:  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD: Putative protein  
; OTHER INFORMATION: sequence of A2058 Autotaxin  
US-08-346-455B-34

Query Match 90.3%; Score 4533; DB 1; Length 829;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 828; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 87 CHDFDELCLKTARGWECTKDRCEVNRNEENACHCEDCLARGDCCTNYQVCKGESHWD 146  
Db 1 CHDFDELCLKTARGWECTKDRCEVNRNEENACHCEDCLARGDCCTNYQVCKGESHWD 60  
QY 147 DCCEIKAAECPAGFVRPPLIIFSVGDGRASYMKKSKVMPNIEKLRSCGTHSPYMRPVY 206  
Db 61 DCCEIKAAECPAGFVRPPLIIFSVGDGRASYMKKSKVMPNIEKLRSCGTHSPYMRPVY 120  
QY 207 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHRWWGGQPLWITATK 266  
Db 121 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHRWWGGQPLWITATK 180  
QY 267 QGVKAGTFFWSVWIPHERRILTLRWLTLPDHERPSVYAFYSEQPDFSGHKYGFPGPEES 326  
Db 181 QGVKAGTFFWSVWIPHERRILTLRWLTLPDHERPSVYAFYSEQPDFSGHKYGFPGPEES 240  
QY 327 SYGSPFTPAKRPRKRVAPKRRQERPAPPKRRRKIHRMDHYAAETRODKMTNPLREIDK 386  
Db 241 SYGSPFTPAKRPRKRVAPKRRQERPAPPKRRRKIHRMDHYAAETRODKMTNPLREIDK 300  
QY 387 IVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNLTNVDDITLVPGTGLGRIR 446  
Db 301 IVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNLTNVDDITLVPGTGLGRIR 360  
QY 447 SKFSNNAKYDPKAIIANLTCKKPDQHFYPYLKQHLPKRLHYANNRRIEDIHLVVERRWHV 506  
Db 361 SKFSNNAKYDPKAIIANLTCKKPDQHFYPYLKQHLPKRLHYANNRRIEDIHLVVERRWHV 420  
QY 507 ARKPLDVYKKPSGKCFQGDHGFNKNVSMQTVFVGYGPTFKYTKVPPFENIELYNVCM 566  
|||||



Db 421 ARKPLDVYKKPSGKCFQGDHGFNDKVNMQTVFVGVGPTFKYKTKVPPFENIELYNVMC 480  
QY 567 DLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEVTRPNYPGIMYLSQDFDLGCTCDDKVE 626  
Db 481 DLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEVTRPNYPGIMYLSQDDDLGCTCDDKVE 540  
QY 627 PKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSY 686  
Db 541 PKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSY 600  
QY 687 TVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCCLAYKNDKQMSYGFLFPPYLSSSSPEAKY 746  
Db 601 TVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCCLAYKNDKQMSYGFLFPPYLSSSSPEAKY 660  
QY 747 DAFLVTNMPMPYPAFKRVWNYFORVLVKKYASERNGVNVISGPIFYDYDGLHDTEDKIK 806  
Db 661 DAFLVTNMPMPYPAFKRVWNYFORVLVKKYASERNGVNVISGPIFYDYDGLHDTEDKIK 720  
QY 807 QYVEGSSIPVPTHYISITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESK 866  
Db 721 QYVEGSSIPVPTHYISITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESK 780  
QY 867 WVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915  
Db 781 WVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 829

RESULT 5

US-08-977-221-34  
; Sequence 34, Application US/08977221  
; Patent No. 6084069  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA; DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/977,221  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: 08/346,455  
; FILING DATE: 28-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149US3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 829

; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE: Melanoma  
; CELL LINE: A2058  
; ORGANELLE:  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: Putative protein  
; OTHER INFORMATION: sequence of A2058 Autotaxin  
US-08-977-221-34

Query Match 90.3%; Score 4533; DB 3; Length 829;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 828; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 87 CHDFDELCLKTARGWECTKDRCGEVRNEENACHCEDCLARGDCCTNYQVVKGESHWVD 146  
Db 1 CHDFDELCLKTARGWECTKDRCGEVRNEENACHCEDCLARGDCCTNYQVVKGESHWVD 60  
QY 147 DDCEEIKAAECFAGFVRPPLIIFSVDFGRASYMKKSGKVMPIEKLRSCTGTHSPMRPVY 206  
Db 61 DDCEEIKAAECFAGFVRPPLIIFSVDFGRASYMKKSGKVMPIEKLRSCTGTHSPMRPVY 120  
QY 207 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHRWGGQPLWITATK 266  
Db 121 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHRWGGQPLWITATK 180  
QY 267 QGVKAGTFFWSVIPHERRILTLRWLTLPDHERPSVAFYSEQPDFSGHKYGFPGPEES 326  
Db 181 QGVKAGTFFWSVIPHERRILTLRWLTLPDHERPSVAFYSEQPDFSGHKYGFPGPEES 240  
QY 327 SYGSPFTPAKRPRKRVAPKRRQERPAPPKRRRKIRHMDHYAAETRODKMTNPLREIDK 386  
Db 241 SYGSPFTPAKRPRKRVAPKRRQERPAPPKRRRKIRHMDHYAAETRODKMTNPLREIDK 300  
QY 387 IVGQMDGLKQLRRCVNVIFVGDHGMEDVTCDRTEFLSNLYTNVDDITLVPGTGRIR 446  
Db 301 IVGQMDGLKQLRRCVNVIFVGDHGMEDVTCDRTEFLSNLYTNVDDITLVPGTGRIR 360  
QY 447 SKFSNNAKYDPKAIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRRIEDHLLVERRWHV 506  
Db 361 SKFSNNAKYDPKAIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRRIEDHLLVERRWHV 420  
QY 507 ARKPLDVYKKPSGKCFQGDHGFNDKVNMQTVFVGVGPTFKYKTKVPPFENIELYNVMC 566  
Db 421 ARKPLDVYKKPSGKCFQGDHGFNDKVNMQTVFVGVGPTFKYKTKVPPFENIELYNVMC 480  
QY 567 DLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEVTRPNYPGIMYLSQDFDLGCTCDDKVE 626  
Db 481 DLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEVTRPNYPGIMYLSQDDDLGCTCDDKVE 540  
QY 627 PKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSY 686  
Db 541 PKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSY 600  
QY 687 TVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCCLAYKNDKQMSYGFLFPPYLSSSSPEAKY 746  
Db 601 TVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCCLAYKNDKQMSYGFLFPPYLSSSSPEAKY 660  
QY 747 DAFLVTNMPMPYPAFKRVWNYFORVLVKKYASERNGVNVISGPIFYDYDGLHDTEDKIK 806

Db 661 DAFLVTNMVPMYPFAKRVWNYFQRLVVKKYASERNGVNVISGPIFDYDGLHDTEDKIK 720  
QY 807 QYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESK 866  
Db 721 QYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESK 780  
QY 867 WVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915  
Db 781 WVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 829

RESULT 6  
PCT-US95-06613-34  
; Sequence 34, Application PC/TUS9506613  
; GENERAL INFORMATION:  
; APPLICANT: STRACKE, MARY; LIOTTA, LANCE;  
; APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,  
; APPLICANT: HENRY; MURATA, JUN  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: 08/346,455  
; FILING DATE: 28-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 829  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: No  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE: Melanoma  
; CELL LINE: A2058  
; ORGANELLE:

; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: Putative protein  
; OTHER INFORMATION: sequence of A2058 Autotaxin  
PCT-US95-06613-34  
  
Query Match 90.3%; Score 4533; DB 5; Length 829;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 828; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 87 CHDFDELCLKTARGWECKDRCEGEVRNEENACHCEDCLARGDCCTNYQVVKGESHWVD 146  
Db 1 CHDFDELCLKTARGWECKDRCEGEVRNEENACHCEDCLARGDCCTNYQVVKGESHWVD 60  
  
QY 147 DCEEIKAAECPAGFVRPPLIIFSVDFRASYSYMKKGVMPNIEKLRSCGTHSPYMRPVY 206  
Db 61 DCEEIKAAECPAGFVRPPLIIFSVDFRASYSYMKKGVMPNIEKLRSCGTHSPYMRPVY 120  
  
QY 207 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHRWGGQPLWITATK 266  
Db 121 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHRWGGQPLWITATK 180  
  
QY 267 QGVKAGTFFWSVIPHERRILTLRWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEES 326  
Db 181 QGVKAGTFFWSVIPHERRILTLRWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEES 240  
  
QY 327 SYGSPFTPAKRPRKRVAPKRRQERPVAPPKRRRKIIHRMDHYAAETRODKMNTNPLREIDK 386  
Db 241 SYGSPFTPAKRPRKRVAPKRRQERPVAPPKRRRKIIHRMDHYAAETRODKMNTNPLREIDK 300  
  
QY 387 IVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIR 446  
Db 301 IVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIR 360  
  
QY 447 SKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLLVERRWHV 506  
Db 361 SKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLLVERRWHV 420  
  
QY 507 ARKPLDVYKKPSGKCFQGDHGFEDNKVNSMQTVFVGYGPTFKYTKVPPFENIELYNVVC 566  
Db 421 ARKPLDVYKKPSGKCFQGDHGFEDNKVNSMQTVFVGYGPTFKYTKVPPFENIELYNVVC 480  
  
QY 567 DLLGLKPAPNNGTHGSLNHLRLNTFRPTMPEEVRPNYPGIMYLGSDFDLGCTCDDKVE 626  
Db 481 DLLGLKPAPNNGTHGSLNHLRLNTFRPTMPEEVRPNYPGIMYLGSDFDLGCTCDDKVE 540  
  
QY 627 PKNKIDELNKRHLTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSY 686  
Db 541 PKNKIDELNKRHLTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSY 600  
  
QY 687 TVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKY 746  
Db 601 TVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKY 660  
  
QY 747 DAFLVTNMVPMYPFAKRVWNYFQRLVVKKYASERNGVNVISGPIFDYDGLHDTEDKIK 806  
Db 661 DAFLVTNMVPMYPFAKRVWNYFQRLVVKKYASERNGVNVISGPIFDYDGLHDTEDKIK 720  
  
QY 807 QYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESK 866  
Db 721 QYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESK 780  
  
QY 867 WVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915  
Db 781 WVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 829

RESULT 7  
US-08-346-455B-67  
; Sequence 67, Application US/08346455B

Patent No. 5731167  
GENERAL INFORMATION:  
APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/346,455B  
FILING DATE: 28-NOV-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 861  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: No  
FEATURE:  
NAME/KEY: N-tera 2D1 ATX protein  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-346-455B-67

Query Match 88.0%; Score 4418; DB 1; Length 861;  
Best Local Similarity 89.1%; Pred. No. 0;  
Matches 822; Conservative 12; Mismatches 19; Indels 70; Gaps 6;  
QY 1 MARRSSFQSCQIIISLFTFAVGVSICLGLFTAHRIKRAEGWEEGPPTVLSDSPWNTISGSCK 60  
DQ 1 MARRSSFQSCQIIISLFTFAVGVSICLGLFTAHRIKRAEGWEEGPPTVLSDSPWNTISGSCK 60  
QY 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECKDRGCEVRNEENACHC 120  
DQ 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECKDRGCEVRNEENACHC 120  
QY 121 SEDCLARGDCCTNYQVCKGESHVWDDCEIEIKAAECFAGFVRPPLIIFSVDFGRASYMK 180  
DQ 121 SEDCLARGDCCTNYQVCKGESHVWDDCEIEIKAAECFAGFVRPPLIIFSVDFGRASYMK 180  
QY 181 KGSKVMPIEKLRSCTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240

Db 181 KGSKVMPIEKLRSCTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240  
QY 241 TFHLRGREKFNHRWGGQPLWTATKQGVKAGTFFWSVVI----PHERRILTILRWLTLP 296  
DQ 241 TFHLRGREKFNHRWGGQPLWTATKQGVKAGTFFWSVVI----PHERRILTILRWLTLP 296  
QY 297 DHER----PSVYAFYSEQDFSGHKYGFPGPEESSYSGSPFTPAKRPKRKVAPKRRQRPV 352  
DQ 297 DHERLRSMPISLSNL-----ISLDTNMPFGPE----- 323  
QY 353 APPKRRRKIHRMDHYAAETRDQKMTNPLREIDKIVGQMDGLKQLKLRRCVNVIFVGDH 412  
DQ 324 -----MTNPLREIDKIVGQMDGLKQLKLRRCVNVIFVGDH 359  
QY 413 GMEDVTCDRTEFLSNLYTNVDDITLVPGTLGRIRSKFSNNAKYDPKAIANLTCKKPDQH 472  
DQ 360 GMEDVTCDRTEFLSNLYTNVDDITLVPGTLG-IRSKFSNNAKYDPKAIANLTCKKPDQH 418  
QY 473 FKPYLKQHLPKRLHYANNRRIEDIHLVVERRHHVARKPLDVYKPSGKCFQGDHGFNDK 532  
DQ 419 FKPYLKQHLPKRLHYANNRRIEDIHLVVERRHHVARKPLDVYKPSGKCFQGDHGFNDK 478  
QY 533 VNSMQTVFVGYGPTFKYKTKVPPFENIELYNMCDLLGLKPAPNNGTHGSLNHLRTNTF 592  
DQ 479 VNSMQTVFVGYGPTFKYKTKVPPFENIELYNMCDLLGLKPAPNNGTHGSLNHLRTNTF 538  
QY 593 RPTMPEEVRPNYPGIMYQSDFDLCTCDDKVEPKNKLDLNLKRLHTKGSTEERHLLYG 652  
DQ 539 RPTMPEEVRPNYPGIMYQSDFDLCTCDDKVEPKNKLDLNLKRLHTKGSTEERHLLYG 598  
QY 653 RPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVS 712  
DQ 599 RPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVS 658  
QY 713 PSFSQNCCLAYKNDKQMSYGFELPPYLSLSSPEAKYDAFLVTNMVPMYPAKRVWNYFQVRL 772  
DQ 659 PSFSQNCCLAYKNDKQMSYGFELPPYLSLSSPEAKYDAFLVTNMVPMYPAKRVWNYFQVRL 718  
QY 773 VKKYASERNGVNIGPIFDYDGLHDTEDKIKQYVEGSIPTVTHYISITSCLDFTQ 832  
DQ 719 VKKYASERNGVNIGPIFDYDGLHDTEDKIKQYVEGSIPTVTHYISITSCLDFTQ 778  
QY 833 PADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVDRDIEHLTSLDFFR 892  
DQ 779 PADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVDRDIEHLTSLDFFR 838  
QY 893 KTSRSYPEILTLKTYLHTYESEI 915  
DQ 839 KTSRSYPEILTLKTYLHTYESEI 861

RESULT 8  
US-08-977-221-67  
Sequence 67, Application US/08977221  
Patent No. 6084069  
GENERAL INFORMATION:  
APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/977,221  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/346,455  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 861  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: No  
FEATURE:  
NAME/KEY: N-tera 2D1 ATX protein  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-977-221-67

Query Match 88.0%; Score 4418; DB 3; Length 861;  
Best Local Similarity 89.1%; Pred. No. 0;  
Matches 822; Conservative 12; Mismatches 19; Indels 70; Gaps 6;

QY 1 MARRSFQSCQIISLFTFAVGVSICLGFTHAHRKRAEGWEEGPPTVLSDSPWNTNIGSCK 60  
Db 1 MARRSFQSCQIISLFTFAVGVSICLGFTHAHRKRAEGWEEGPPTVLSDSPWNTNIGSCK 60  
QY 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECKDRCEVRNEENACHC 120  
Db 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECKDRCEVRNEENACHC 120  
QY 121 SEDCLARGDCCTNYQVCKGESHVDDDCIEIKAAECPAGFVRPPLIIFSVGDFRASYMK 180  
Db 121 SEDCLARGDCCTNYQVCKGESHVDDDCIEIKAAECPAGFVRPPLIIFSVGDFRASYMK 180  
QY 181 KGSVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240  
Db 181 KGSVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240  
QY 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVI----PHERILTLRWLTLP 296  
Db 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVI----PHERILTLRWLTLP 296  
QY 297 DHER-----PSVYAFYSEQPDFSGHKYGFPGPEESSYSGSPFPAKRPKRKVAPKRRQRPV 352  
Db 297 DHERLRSMPSILSNL-----ISLDTNMPFGPE----- 323  
QY 353 APPKRRRKIHRMDHYAAETRQDKMTNPLREIDKIVGQMDGLKQLKLRRCVNVIFVGDH 412  
Db 324 -----MTNPLREIDKIVGQMDGLKQLKLRRCVNVIFVGDH 359  
QY 413 GMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYDPKAILANLTCKKPDQH 472  
Db 360 GMEDVTCDRTEFLSNYLTNVDDITLVPGTLG-IRSKFSNNAKYDPKAILANLTCKKPDQH 418

QY 473 FKPYLKQHLPKRLHYANNRRRIEDIHLLVERRHWVARKPLDVYKKPSGKCFQGDHGFNDK 532  
Db 419 FKPYLKQHLPKRLHYANNRRRIEDIHLLVERRHWVARKPLDVYKKPSGKCFQGDHGFNDK 478  
QY 533 VNSMQTVFVGYPGTFKTKVPPFENIELYNVMCDLGLKPAPNNGTHGSLNHLRTNTF 592  
Db 479 VNSMQTVFVGYPGTFKTKVPPFENIELYNVMCDLGLKPAPNNGTHGSLNHLRTNTF 538  
QY 593 RPTMPEEVTRPNYPGIMYQSDFDLCTCDDKVEPKNKLDLNLKRLHTKKGSTEERHLLYG 652  
Db 539 RPTMPEEVTRPNYPGIMYQSDFDLCTCDDKVEPKNKLDLNLKRLHTKKGSTEERHLLYG 598  
QY 653 RPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVS 712  
Db 599 RPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVS 658  
QY 713 PSFSQNCCLAYKNDKQMSYGFPPYLSSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRVL 772  
Db 659 PSFSQNCCLAYKNDKQMSYGFPPYLSSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRVL 718  
QY 773 VKKYASERNGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQ 832  
Db 719 VKKYASERNGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQ 778  
QY 833 PADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFR 892  
Db 779 PADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFR 838  
QY 893 KTSRYPEILTLKTYLHTESEI 915  
Db 839 KTSRYPEILTLKTYLHTESEI 861

RESULT 9  
PCT-US95-06613-67  
; Sequence 67, Application PC/TUS9506613  
; GENERAL INFORMATION:  
; APPLICANT: STRACKE, MARY; LIOTTA, LANCE;  
; APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,  
; APPLICANT: HENRY; MURATA, JUN  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/346,455  
; FILING DATE: 28-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434

REFERENCE/DOCKET NUMBER: 2026-4149US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 861  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: No  
FEATURE:  
NAME/KEY: N-tera 2D1 ATX protein  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PCT-US95-06613-67

Query Match 88.0%; Score 4418; DB 5; Length 861;  
Best Local Similarity 89.1%; Pred. No. 0;  
Matches 822; Conservative 12; Mismatches 19; Indels 70; Gaps 6;

QY 1 MARRSSFQSQIISLFFAVGVSIICLFTAHRRIKRAEGWEEGPPPTVLSDSPWTNIGSCK 60  
Db 1 MARRSSFQSQIISLFFAVGVNICLFTAHRRIKRAEGWEEGPPPTVLSDSPWTNIGSCK 60  
QY 61 GRCFELQEAQPPDCRCNLCCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHC 120  
Db 61 GRCFELQEAQPPDCRCNLCCKSYTSCCHDFDELCLKTARAWECTKDRCGEVRNEENACHC 120  
QY 121 SEDCLARGDCCTNYQVYKGGESHVDDDCCEEIKAAECFAGFVRPPLIIFSVDFRASVMK 180  
Db 121 SEDCLARGDCCTNYQVYKGGESHVDDDCCEEIKAAECFAGFVRPPLIIFSVDFRASVMK 180  
QY 181 KSKVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVDA 240  
Db 181 KSKVMPNIEKLRSCGTHSPHMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVDA 240  
QY 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFNSVVI- - - - -PHERRILTLRLWLTLP 296  
Db 241 TFHLRGREKFNHRWGGQPLWITATKQGES- - - - -WNILLVCCHPRAEILTILQWLTL 296  
QY 297 DHER- - - - -PSVYAFYSEQPDFSGHKYGPFGPESSYSGSPFPAKRPKRKVAPKRRQRPV 352  
Db 297 DHERLRSMPSILSNL- - - - -ISLDTNMPFGPE- - - - - 323  
QY 353 APPKRRRKIHRMDHYAAETRDQKMTNPLREIDKIVGQMDGLKQLKLRRCVNVIFVGDH 412  
Db 324 - - - - -MTNPLREIDKIVGQMDGLKQLKLRRCVNVIFVGDH 359  
QY 413 GMEDVTCDRTEFLSNVLTNVDDITLVPGLTGRIRSFNNKAYDPKAIIANLTCKKPDQH 472  
Db 360 GMEDVTCDRTEFLSNVLTNVDDITLVPGLTGRIRSFNNKAYDPKAIIANLTCKKPDQH 418  
QY 473 FKPYLKQHLPKRLHYANNRRIEDIHLVRRWHVARKPLDVYKKPSGCKFFQGDHGFNDK 532  
Db 419 FKPYLKQHLPKRLHYANNRRIEDIHLVRRWHVARKPLDVYKKPSGCKFFQGDHGFNDK 478  
QY 533 VNSMQTVFVGYPFKYKTKVPPFENIELYNVCMCDLGLKPAENNGTHGSLNHLRLTNTF 592  
Db 479 VNSMQTVFVGYPFKYKTKVPPFENIELYNVCMCDLGLKPAENNGTHGSLNHLRLTNTF 538  
QY 593 RPTMPEEVRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDELNKLRLHTKGSTEEHLLYG 652  
Db 539 RPTMPEEVRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDELNKLRLHTKGSTEEHLLYG 598  
QY 653 RPAVLYRTRYDILYHTDFESGYSEIFLMLLTSTVSKQAEVSSVPDHLTSCVRPDVRVS 712  
Db 599 RPAVLYRTRYDILYHTDFESGYSEIFLMLLTSTVSKQAEVSSVPDHLTSCVRPDVRVS 658  
QY 713 PSFSQNCILAYKNDKQMSYGFLEPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRLV 772

Db 659 PSFSQNCILAYKNDKQMSYGFLEPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRLV 718  
QY 773 VKKYASERNGVNVISGPIFDYDGLDHTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQ 832  
Db 719 VKKYASERNGVNVISGPIFDYDGLDHTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQ 778  
QY 833 PADKCDGPLSVSFFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFR 892  
Db 779 PADKCDGPLSVSFFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFR 838  
QY 893 KTSRYPEILTLKTYLHTYESEI 915  
Db 839 KTSRYPEILTLKTYLHTYESEI 861  
RESULT 10  
US-08-346-455B-38  
; Sequence 38, Application US/08346455B  
; Patent No. 5731167  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA; DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,455B  
; FILING DATE: 28-NOV-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 979  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: No  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:

```

; TISSUE TYPE: Liver
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; OTHER INFORMATION:
US-08-346-455B-38

```

Query Match	85.3%;	Score 4279;	DB 1;	Length 979;
Best Local Similarity	87.1%;	Pred. NO. 0;		
Matches 805;	Conservative 13;	Mismatches 36;	Indels 70;	Gaps 8;

QY	1	MARRSSFQSCQIIISLFTFAVGVSI	CLGFTAHRIKRAEGWEEGPTVLSDSPWTNISG	60	
Db	1	MARRSSFQSCQDISLFTFAVGVNICL	GLGFTAHRIKRAEGWEEGPTVLSDSPWTNISG	60	
QY	61	GRCFELQEA	PPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDR	CEVVRNEENACHC 120	
Db	61	GRCFELQEA	PPDCRCDNLCKSYTSCCHDFDELCLKTARAWECTKDR	CEVVRNEENACHC 120	
QY	121	SEDC	LARGDCCTNYQVVKKGESHVWDDCEIEIKAAECPAGFVR	PLIIFSV	DGF--RASY 178
Db	121	SEDC	LARGDCCTNYQVVKKGESHVWDDCEIEIKAAECLQ--VC	SPSINHLLRGLWLPMTSY 178	
QY	179	MKKGSKVMPNIE	KRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVG	NSMYDPVF 238	
Db	179	MKKGSKVMPNIE	KRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVG	NSMYDPVF 238	
QY	239	DATFHLRGREKFN	RHWGGQPLWITATKQGVKAGTFFWSVVI---PHERRIL	TILRWLT 294	
Db	239	DATFHLRGREKFN	RHWGGQPLWITATKORGES---WNILLVCCHPSRAEILT	ILQWLT 294	
QY	295	LPDHERPSVYAFY	SEQDFSGHKYGPFGPEESSYSGSPFTPAKRPRK	KVAPKRQERPAP 354	
Db	295	LPDHERPSVYAFY	SEQDFSGHKHMPFGPE-----	324	
QY	355	PKRRRKIH	RMDHYAAETRODKMTNPLREIDKIVGQLMDGLKQLKLR	RCVNVIVFGDHGM 414	
Db	325	-----	-----MTNPLREMHKIVGQLMDGLKQLKLR	CVNVIVFV---ET_359	
QY	415	EDVTCD--	RTEFLSNLYTNVDDITLVPGLGRIRSKFSNNAKYDP	KAIANLTC	KKPDQH 472
Db	360	MDGRCHMYR	TEFLSNLYTNVDDITLVPGLGRIRSKFSNNAKYDP	KAIANLTC	KKPDQH 419
QY	473	FKPYLKQHL	PKRLHYANNRRITEDIHLLVERRRWHVARKPLDVYK	PSGCKFFQGDHGF	DNK 532
Db	420	FKPYLKQHL	PKRLHYANNRRITEDIHLLVERRRWHVARKPLDVYK	PSGNAFSRETTAF	DNK 479
QY	533	VNSMQTVFVG	YGPTFKYKTKVPPFENIELYNVCMDDLGLKPA	PNNGTGSLNHL	LRTNTF 592
Db	480	VNSMQTVFVG	YGPTFKYKTKVPPFENIELYNVCMDDLGLKPA	PNNGTGSLNHL	LRTNTF 539
QY	593	RPTMPEEV	TRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLD	ELNKR	LHTKGSTEERHLLYG 652
Db	540	RPTMPEEV	TRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLD	ELNKR	LHTKGSTEERHLLYG 599
QY	653	-RPAVLYR	TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEV	SSVPDHLTSCVR	PDVRV 711
Db	600	DRPAVLYR	TRYDILYHTDFESGYSEIFLMLPLWTSYTVSKQAEV	SSVPDHLTSCVR	PDVRV 659
QY	712	SPSFSQNC	LAYKNDKQMSYGFLEPPYLSSSPEAKYDAFLVTN	MPMYPAFKRV	VWNYFQRV 771
Db	660	SPSFSQNC	LAYKNDKQMSYGFLEPPYLSSSPEAKYDAFLVTN	MPMYPAFKRV	VWNYFQRV 719
QY	772	LVKKYAS	ERNGVNVISGPIFDYDYGDLHDTEDKIKQYVEG	SSIPVPTHYYS	IITSCLDFT 831
Db	720	LVKKYAS	ERNGVNVISGPIFDYDYGDLHDTEDKIKQYVEG	SSIPVPTHYYS	IITSCLDFT 779



IDENTIFICATION METHOD:  
OTHER INFORMATION: putative autotaxin  
OTHER INFORMATION: protein sequence from human liver  
US-08-977-221-38

Query Match 85.3%; Score 4279; DB 3; Length 979;  
Best Local Similarity 87.1%; Pred. No. 0;  
Matches 805; Conservative 13; Mismatches 36; Indels 70; Gaps 8;

QY	1	MARRSSFQSCQIISLFTFAVGVSICLGF	TAHRIKRAEGWEEGPP	TVLSDSPWTN	ISGSCK	60
Db	1	MARRSSFQSCQDISLFTFAVGVNICLGF	TAHRIKRAEGWEEGPP	TVLSDSPWTN	ISGSCK	60
QY	61	GRCFELQEA	PPDCRCDNLCKSYTSCCHDFE	DELCLKTARGWECTK	DRCGEV	NEENACHC 120
Db	61	GRCFELQEA	PPDCRCDNLCKSYTSCCHDFE	DELCLKTARAWECTK	DRCGEV	NEENACHC 120
QY	121	SEDCLARGDCCTNYQVCKGESHVDD	CCEIKAAECPAGFVRPPL	IIFSVDGF--	RASY	178
Db	121	SEDCLARGDCCTNYQVCKGESHVDD	CCEIKAAECLQ--VCSP	SINHLRLRGWLP	MTSY	178
QY	179	MKKGSKVMPNIEKLRSCGTHSPYMRPV	YPTKTFPNLYTLATGLYPESHG	IVGNSMYDPVF		238
Db	179	MKKGSKVMPNIEKLRSCGTHSPYMRPV	YPTKTFPNLYTLATGLYPESHG	IVGNSMYDPVF		238
QY	239	DATFHLRGREKFNHRWGGQPLWITATK	QGVKAGTFEWSVVI----	PHERRIL	TILRWLT	294
Db	239	DATFHLRGREKFNHRWGGQPLWITATK	QGES-----WNILL	VCCHPSRAE	ILTLOWLT	294
QY	295	LPDHERPSVYAFYSEQPDFSGHKYGP	FGPESSYCSGSPFTPAKRPKR	KVAPKRR	QERP	VAP 354
Db	295	LPDHERPSVYAFYSEQPDFSGHKHMP	FGPE-----			324
QY	355	PKKRRRKIHRMDHYAAETQDKMTNPL	NEIDKIVGQLMDGLKQLRRCV	NVIFVGDHGM		414
Db	325	-----	-----MTNPL	REMCHKIVGQLMDGLKQLRRCV	NVIFV-----	ET 359
QY	415	EDVTCD--RTEFLSNVLTNVDDITL	VPGLGRIRSKFSNNAKYDPKAI	IANLTCK	KPDQH	472
Db	360	MDGRCHMYRTEFLSNVLTNVDDITL	VPGLGRIRSKFSNNAKYDPKAI	IANLTCK	KPDQH	419
QY	473	FKPYLKQHLPKRLHYANNRRIEDIHL	LVERRVHARKPLDVYKKPSG	KCFQGDHGD	DNK	532
Db	420	FKPYLKQHLPKRLHYANNRRIEDIHL	LVERRVHARKPLDVYKKPSG	NAFSRET	TAEDNK	479
QY	533	VNSMQTVFVGPTFKYKTKVPPFENIE	LYNVMDLLGLKPAPNNGTHG	SLNHL	LRNTF	592
Db	480	VNSMQTVFVGPTFKYKTKVPPFENIE	LYNVMDLLGLKPAPNNGTHG	SLNHL	LRNTF	539
QY	593	RPTMPEEVRPNYPGIMYLSDFDLG	CTCDDKVEPKNKLDLNLKRL	HTKGST	EEHLLYG	652
Db	540	RPTMPEEVRPNYPGIMYLSDFDLG	CTCDDKVEPKNKLDLNLKRL	HTKGST	EEHLLYG	599
QY	653	-RPAVLYRTRVDILYHTDFESGYSE	IFLMLLWTSYTVSKQAEVSS	VPDHLT	SCVRPD	VRV 711
Db	600	DRPAVLYRTRVDILYHTDFESGYSE	IFLMLWTSYTVSKQAEVSS	VPDHLT	SCVRPD	VRV 659
QY	712	SPFSQNC	LAYKNDKQMSYGFLPPYLSS	SSPEAKYDAFLVTNMV	MPYPAF	KRVWNYFQRV 771
Db	660	SPFSQNC	LAYKNDKQMSYGFLPPYLSS	SSPEAKYDAFLVTNMV	MPYPAF	KRVWNYFQRV 719
QY	772	LVKKYASERN	GVNVISGPIFDYDYG	LHDTEDKIKQYVEG	SSIPVPTHY	SIITSCLDFT 831
Db	720	LVKKYASERN	GVNVISGPIFDYDYG	LHDTEDKIKQYVEG	SSIPVPTHY	SIITSCLDFT 779
QY	832	QPADKCDG	PLSVSSFILPHRPDNEESC	NSSDESKWVEELMK	MHTARVRD	IEHLTSLDFF 891
Db	780	QPADKCDG	PLSVSSFILPHRPDNEESC	NSSDESKWVEELMK	MHTARVRD	IEHLTSLDFF 839
QY	892	RKTSR	SYPEILTLKTYLHTYESEI		915	
Db	840	RKTSR	SYPEILTLKTYLHTYESEI		863	

RESULT 12  
PCT-US95-06613-38  
; Sequence 38, Application PC/TUS9506613  
; GENERAL INFORMATION:  
; APPLICANT: STRACKE, MARY; LIOTTA, LANCE;  
; APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,  
; APPLICANT: HENRY; MURATA, JUN  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/346,455  
; FILING DATE: 28-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 979  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: No  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE: Liver  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: putative autotaxin  
; OTHER INFORMATION: protein sequence from human liver  
PCT-US95-06613-38

Query Match 85.3%; Score 4279; DB 5; Length 979;

Best Local Similarity 87.1%; Pred. No. 0;		Matches 805; Conservative 13; Mismatches 36; Indels 70; Gaps 8;	
QY	1	MARRSSFQSCQIIISLFFFAVGVSICLGFTAHRIKRAEGWEGPPTVLSDPWTNISGSK	60
Db	1	MARRSSFQSCQDISLFFFAVGVSICLGFTAHRIKRAEGWEGPPTVLSDPWTNISGSK	60
QY	61	GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCKLTARGWECTKDRCGEVNRNEENACHC	120
Db	61	GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCKLTARAWECTKDRCGEVNRNEENACHC	120
QY	121	SEDCLARGDCCTNYQVCKGESHVWDDCEEEIKAAECAPGVRPPPLIIFSVDFG--RASY	178
Db	121	SEDCLARGDCCTNYQVCKGESHVWDDCEEEIKAAECQ--VCSPSINHLRLGWLPMTSY	178
QY	179	MKKGSKVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVF	238
Db	179	MKKGSKVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVF	238
QY	239	DATFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVI----PHERRILTLRWLT	294
Db	239	DATFHLRGREKFNHRWGGQPLWITATKQGES----WNILVCHPSRAEILTILQWLT	294
QY	295	LPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYSGSPFTPAKRPKRKVAPKRQERPAP	354
Db	295	LPDHERPSVYAFYSEQPDFSGHKMPFGE-----	324
QY	355	PKRRRKIHRMDHYAAETRODKMNTNPLREIDKIVGQMLDGLKQLKLRRCVNVIFVGDHGM	414
Db	325	-----MTNPLREMHKIVGQMLDGLKQLKLRRCVNVIFV---ET	359
QY	415	EDVTCD--RTEFLSNLYLTNVDDITLVPGLTGLRIRSKFSNNAKYDPAKAIANLTCKKPDQH	472
Db	360	MDGRCHMYRTEFLSNLYLTNVDDITLVPGLTGLRIRSKFSNNAKYDPAKAIANLTCKKPDQH	419
QY	473	FKPYLKQHLPKRLHYANNRRIEDIHLVERRRWHVARKPLDVYKPSGKCFQGDHGFNDK	532
Db	420	FKPYLKQHLPKRLHYANNRRIEDIHLVERRRWHVARKPLDVYKPSGNAFSRETTAFDNK	479
QY	533	VNSMQTVFVGYGPTFKYKTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRLTNTF	592
Db	480	VNSMQTVFVGYGPTFKYKTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRLTNTF	539
QY	593	RPTMPEEVRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEERHLLYG	652
Db	540	RPTMPEEVRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEERHLLYG	599
QY	653	-RPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRV	711
Db	600	DRPAVLYRTRYDILYHTDFESGYSEIFLMPPLWTSYTVSKQAEVSSVPDHLTSCVRPDVRV	659
QY	712	SPSFQONCLAYKNDQMSYGFPPYLSSSSPEAKYDAFLVTNMVMPYPAFKRVWNYFQRV	771
Db	660	SPSFQONCLAYKNDQMSYGFPPYLSSSSPEAKYDAFLVTNMVMPYPAFKRVWNYFQRV	719
QY	772	LVKKYASERNGVNVISGPIFDYDYGDLHDTEDKIKQYVEGSSIPVPTHYYSIITSLDFT	831
Db	720	LVKKYASERNGVNVISGPIFDYDYGDLHDTEDKIKQYVEGSSIPVPTHYYSIITSLDFT	779
QY	832	QPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFF	891
Db	780	QPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFF	839
QY	892	RKTSRSPYEILTTLKTYLHYEISEI	915
Db	840	RKTSRSPYEILTTLKTYLHYEISEI	863

RESULT 13  
US-08-346-455B-36  
; Sequence 36, Application US/08346455B  
; Patent No. 5731167  
; GENERAL INFORMATION:

APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/346,455B  
FILING DATE: 28-NOV-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 788  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: No  
ORIGINAL SOURCE:  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE: teratocarcinoma  
CELL LINE: N-tera 2D1  
ORGANELLE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: N-tera 2D1 putative  
OTHER INFORMATION: ATX protein sequence  
US-08-346-455B-36

Query Match . 76.5%; Score 3842; DB 1; Length 788;  
Best Local Similarity 85.6%; Pred. No. 0;  
Matches 727; Conservative 12; Mismatches 40; Indels 70; Gaps 8;

QY 76 CDNLCKSYTSCCHDFDELCKLTARGWECTKDRCGEVNRNEENACHCSEDCLARGDCCTNYQ 135  
Db 1 CDNLCKSYTSCCHDFDELCKLTARAWECTKDRCGEVNRNEENACHCSEDCLARGDCCTNYQ 60

QY 136 VVCKGESHVDDDCCEEIKAAECFAGFVRPPLIIFSVDF--RASYMKKSGSKVMPNIEKLR 193  
Db 61 VVCKGESHVDDDCCEEIKAAECFAGFVRPPLIIFSVDF--RASYMKKSGSKVMPNIEKLR 118  
QY 194 SCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHR 253  
Db 119 SCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHR 178  
QY 254 WVGQPLWITATKQGVKAGTFWVVI---PHERRILTLRLWTLPLDHERPSVYAFYSE 309  
Db 179 WVGQPLWITATKQGVKAGTFWVVI---PHERRILTLRLWTLPLDHERPSVYAFYSE 234  
QY 310 QPDFSGHKYGPFGPESSYSGSPFTPAKRPKRKVAPKRRQRPVAPPKRRRKIHRMDHYA 369  
Db 235 QPDFSGHKHMPFGPE----- 249  
QY 370 AETRDQKMTNPLREIDKIVGQMDGLKQLKRRRCVNVIFVGDHGMEDVTCD--RTEFLSN 427  
Db 250 -----MPNPLREMHKIVGQMDGLKQLKLRVCNVVIFV---ETMDGRCHMYRTEFLSN 299  
QY 428 YLTNVDDITLVPGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHY 487  
Db 300 YLTNVDDITLVPGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHY 359  
QY 488 ANNRRIEDIHLLVERRWHVARKPLDVYKKPSGCKCFQGDHGDKNVNSMQTVFVGYPGF 547  
Db 360 ANNRRIEDIHLLVERRWHVARKPLDVYKKPSGNAFSAFRTTAADKNVNSMQTVFVGYPGF 419  
QY 548 KYTKVPPFENIELYNVMDLLGLKPAAPNNGTHGSLNHLRLTNTFRPTMPEEVRPNYPG 607  
Db 420 KYTKVPPFENIELYNVMDLLGLKPAAPNNGTHGSLNHLRLTNTFRPTMPEEVRPNYPG 479  
QY 608 IMYLSDFDLGCTCDDKVEPKNKIDELNKLRLTKGSTEERHLLYG-RPAVLYRTRYDILY 666  
Db 480 IMYLSDFDLGCTCDDKVEPKNKIDELNKLRLTKGSTEERHLLYGDRPAVLYRTRYDILY 539  
QY 667 HTDFESGYSEIFLMLLTSTYVSKQAEVSSVPDHLTSCVRPDVRSVSPFSQNCCLAYKNDK 726  
Db 540 HTDFESGYSEIFLMLLTSTYVSKQAEVSSVPDHLTSCVRPDVRSVSPFSQNCCLAYKNDK 599  
QY 727 QMSYGFLLFPYLSSSPEAKYDAFLVTNMVMPYPAFKRVWNYFQRLVKKYASERNGVNVI 786  
Db 600 QMSYGGGLPPYLSSSPEAKYDAFLVTNMVMPYPAFKRVWNYFQRLVKKYASERNGVNVI 659  
QY 787 SGPIFDYDGLHDTEDTKIQYVEGSSIPVPHYYSIITSCLDFTQPADKCDGPLSVSSF 846  
Db 660 SGPIFDYDGLHDTEDTKIQYVEGSSIPVPHYYSIITSCLDFTQPADKCDGPLSVSSF 719  
QY 847 ILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSSRYPEILTLKT 906  
Db 720 ILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSSRYPEILTLKT 779  
QY 907 YLHTYESEI 915  
Db 780 YLHTYESEI 788

RESULT 14  
US-08-977-221-36  
; Sequence 36, Application US/08977221  
; Patent No. 6084069  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA; DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK

; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/977,221  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/346,455  
; FILING DATE: 28-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149US3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 788  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE: teratocarcinoma  
; CELL LINE: N-tera 2D1  
; ORGANELLE:  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: N-tera 2D1 putative  
; OTHER INFORMATION: ATX protein sequence  
US-08-977-221-36

Query Match 76.5%; Score 3842; DB 3; Length 788;  
Best Local Similarity 85.6%; Pred. No. 0;  
Matches 727; Conservative 12; Mismatches 40; Indels 70; Gaps 8;

QY 76 CDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCEDCLARGDCCTNYQ 135  
Db 1 CDNLCKSYTSCCHDFDELCLKTARAWECTKDRCGEVRNEENACHCEDCLARGDCCTNYQ 60  
QY 136 VVCKGESHVDDDCCEEIKAAECFAGFVRPPLIIFSVDF--RASYMKKSGSKVMPNIEKLR 193  
Db 61 VVCKGESHVDDDCCEEIKAAECFAGFVRPPLIIFSVDF--RASYMKKSGSKVMPNIEKLR 118  
QY 194 SCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHR 253  
Db 119 SCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHR 178  
QY 254 WVGQPLWITATKQGVKAGTFWVVI---PHERRILTLRLWTLPLDHERPSVYAFYSE 309  
Db 179 WVGQPLWITATKQGVKAGTFWVVI---PHERRILTLRLWTLPLDHERPSVYAFYSE 234





Db 300 YLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHY 359

QY 488 ANNRRIEDIHLLVERWHVARKPLDVYKKPSGKCFQGDHGFNDKNVSMQTVFVGYGPTF 547  
|||||

Db 360 ANNRRIEDIHLLVERWHVARKPLDVYKKPSGNAFSRETAFDNKNVSMQTVFVGYGPTF 419  
|||||

QY 548 KYTKVPPFENIELYNVWCDLLGLKPAPNNGTHGSLNHLRTNTFRPTMPEEVRPNYPG 607  
|||||

Db 420 KYTKVPPFENIELYNVWCDLLGLKPAPNNGTHFSLNHLRTNTFRPTMPEEVRPNYPG 479  
|||||

QY 608 IMYLSDFDLGCTCDKVEPKNKLDLKNKRLHTKSGTEERHLLYG - RPAVLYRTRYDILY 666  
|||||

Db 480 IMYLSDFDLGCTCDKVEPKNKLDLKNKRLHTKSGTEERHLLYGDRPAVLYRTRYDILY 539  
|||||

QY 667 HTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRSVSPFSQNCLAYKNDK 726  
|||||

Db 540 HTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRSVSPFSQNCLAYKNDK 599  
|||||

QY 727 QMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRVLVKKYASERNGVNVI 786  
|||||

Db 600 QMSYGGLGPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRVLVKKYASERNGVNVI 659  
|||||

QY 787 SGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSF 846  
|||||

Db 660 SGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSF 719  
|||||

QY 847 ILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKT 906  
|||||

Db 720 ILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKT 779  
|||||

QY 907 YLHTYESEI 915  
|||||

Db 780 YLHTYESEI 788

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2001, 14:43:22 ; Search time 25.97 Seconds  
(without alignments)  
2683.855 Million cell updates/sec

Title: US-09-483-831-69  
Perfect score: 5019  
Sequence: 1 MARRSSFQSQIISLFTFAV.....RSYPEILTLKTYLHTYESEI 915

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5019	100.0	915	1 A55144	autotaxin precursor
2	4306	85.8	885	1 A55453	plasma cell membra
3	2084	41.5	875	1 A57080	cell surface antig
4	1955.5	39.0	925	1 A39216	plasma cell membra
5	1899.5	37.8	905	1 A27410	plasma cell membra
6	646.5	12.9	457	2 T09932	nucleotide pyropho
7	612	12.2	496	2 T09931	nucleotide pyropho
8	611.5	12.2	479	2 T03293	nucleotide pyropho
9	610	12.2	829	2 T19494	hypothetical prote
10	609	12.1	461	2 T09933	nucleotide pyropho
11	539.5	10.7	485	2 T40657	probable phophodie
12	522	10.4	429	2 T33724	hypothetical prote
13	461	9.2	433	2 B82537	phosphodiesterase-
14	456.5	9.1	674	2 T19495	hypothetical prote
15	456	9.1	300	2 A41179	protein kinase PC-
16	410	8.2	743	2 S19437	hypothetical prote
17	335.5	6.7	493	2 S50443	hypothetical prote
18	297.5	5.9	614	2 T30973	hypothetical prote
19	241.5	4.8	133	2 T09934	hypothetical prote
20	238	4.7	96	2 A25274	phosphodiesterase
21	238	4.7	453	2 T16795	hypothetical prote
22	163	3.2	360	2 T20867	hypothetical prote
23	146.5	2.9	329	1 NCBYN1	nuclease NUC1 (EC
24	145.5	2.9	465	2 T34936	hypothetical prote
25	142	2.8	1186	2 T03180	tyrosine protein k
26	132	2.6	462	2 T36185	hypothetical prote
27	127	2.5	1820	2 A55494	latent transformin
28	123.5	2.5	454	2 A46498	glucocorticoid-sen
29	122	2.4	1374	2 S70712	protein-tyrosine k

30	121.5	2.4	2471	2 T03820	probable histidine
31	121	2.4	1888	2 T39009	hypothetical prote
32	119.5	2.4	475	2 A38340	66K glycoprotein p
33	119	2.4	2019	1 JQ1322	tenascin precursor
34	117	2.3	335	2 T11588	nuclease NUC1 (EC
35	116.5	2.3	989	2 S69711	hypothetical prote
36	112	2.2	1035	2 S61342	sulfite reductase
37	112	2.2	1394	2 A35626	transforming growt
38	112	2.2	2911	2 T20566	hypothetical prote
39	111	2.2	1394	2 A29637	position-specific
40	111	2.2	3002	2 A47221	fibrillin 1 precur
41	110.5	2.2	2871	2 A55567	fibrillin I - bovi
42	110.5	2.2	3147	2 T18674	hypothetical prote
43	110	2.2	753	2 T28787	hypothetical prote
44	109.5	2.2	476	1 SGMSV	vitronectin precur
45	109.5	2.2	1472	2 S67195	probable membrane

ALIGNMENTS

RESULT 1  
A55144  
autotaxin precursor - human  
N;Contains: phosphodiesterase I (EC 3.1.4.1)  
C;Species: Homo sapiens (man)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A55144; A42329  
R;Murata, J.; Lee, H.Y.; Clair, T.; Krutzsch, H.C.; Arestad, A.A.; Sobel, M.E.; Liott  
J. Biol. Chem. 269, 30479-30484, 1994  
A;Title: cDNA cloning of the human tumor motility-stimulating protein, autotaxin, rev  
A;Reference number: A55144; MUID:95074054  
A;Accession: A55144  
A;Molecule type: mRNA  
A;Residues: 1-915 <MUR>  
A;Cross-references: GB:L35594; NID:g537905; PIDN:AAA64785.1; PID:g537906  
A;Note: parts of this sequence were confirmed by peptide sequencing  
R;Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.; Cioce, V.; Schiffmann,  
J. Biol. Chem. 267, 2524-2529, 1992  
A;Title: Identification, purification, and partial sequence analysis of autotaxin, a  
A;Reference number: A42329; MUID:92129337  
A;Accession: A42329  
A;Molecule type: protein  
A;Residues: 256-266;422-444;504-507,'AN';510,'X',511-515;533-548;'S',554-559,'N',561-  
A;Experimental source: A2058 melanoma cells  
A;Note: sequence extracted from NCBI backbone (NCBIP:78526, NCBIP:78523, NCBIP:78521,  
A;Note: a peptide fragment Tyr-Asp-Val-Pro-Trp-Asn-Glu-Thr-Ile was also found  
C;Comment: This protein acts as an autocrine factor to stimulate tumor cell motility.  
C;Genetics:  
A;Gene: GDB:ATX  
A;Cross-references: GDB:378346  
A;Map position: 8q22-8qter  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: EF hand; glycoprotein; phosphoprotein; phosphoric diester hydrolase  
F;55-98/Domain: somatomedin B homology <SBH1>  
F;99-142/Domain: somatomedin B homology <SBH2>  
F;54,463,577,859/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;210/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 100.0%; Score 5019; DB 1; Length 915;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MARRSSFQSQIISLFTFAVGVSI	CLGFTAHRIKRAEGWEEGPTVLSDSPWTNISGSCK	60
Db	1	MARRSSFQSQIISLFTFAVGVSI	CLGFTAHRIKRAEGWEEGPTVLSDSPWTNISGSCK	60
QY	61	GRCFELQEA	AGPPDCRCDNLCKSYTSCCHDFDELCKLTARGWECTKDRCGEVRNEENACHC	120
Db	61	GRCFELQEA	AGPPDCRCDNLCKSYTSCCHDFDELCKLTARGWECTKDRCGEVRNEENACHC	120
QY	121	SEDC	LARGDCCTNYQVVCKGESHVDDDC	EEIKAAEC
			PAGFVRPPLIIFSVDGFRASYMK	180



Db	121	SEDCLARGDCCTNYQVVKGESHVWDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMK	180
QY	181	KGSKVMPNIEKLKSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA	240
Db	181	KGSKVMPNIEKLKSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA	240
QY	241	TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPHERRLTILRWLTLPDHER	300
Db	241	TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPHERRLTILRWLTLPDHER	300
QY	301	PSVYAFYSEQPDFSGHKYGPFGPESSYSGSPFPAKRPKRKVAPKRQERPVPAPPKRRR	360
Db	301	PSVYAFYSEQPDFSGHKYGPFGPESSYSGSPFPAKRPKRKVAPKRQERPVPAPPKRRR	360
QY	361	KIHRMDHYAAETRODKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTC	420
Db	361	KIHRMDHYAAETRODKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTC	420
QY	421	RTEFLSNLYLTNVDDITLVPGTLGRIRSKFSNNAKYDPKAIIANLTCKPDQHEKPYLKQH	480
Db	421	RTEFLSNLYLTNVDDITLVPGTLGRIRSKFSNNAKYDPKAIIANLTCKPDQHEKPYLKQH	480
QY	481	LPKRLHYANNRRIEDIHLVRRWHVARKPLDVYKKPSGKCFQDGHGFDNKNVSMQTVF	540
Db	481	LPKRLHYANNRRIEDIHLVRRWHVARKPLDVYKKPSGKCFQDGHGFDNKNVSMQTVF	540
QY	541	VGYGPTFKYKTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRTNTFRPTMPEEV	600
Db	541	VGYGPTFKYKTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRTNTFRPTMPEEV	600
QY	601	TRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDLNLKRLHTKSGSTERHLLYGRPAVLRT	660
Db	601	TRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDLNLKRLHTKSGSTERHLLYGRPAVLRT	660
QY	661	RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCL	720
Db	661	RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCL	720
QY	721	AYKNDKQMSYGFLFPPYLSSSSPEAKYDAFLVTNMVPMYPAFKRVWNVFQRLVKKYASER	780
Db	721	AYKNDKQMSYGFLFPPYLSSSSPEAKYDAFLVTNMVPMYPAFKRVWNVFQRLVKKYASER	780
QY	781	NGVNVISGPIFDYDYGDLHDTEDKIKQYVEGSSIPVPTHYYSIITSCILDFTQPADKCDGP	840
Db	781	NGVNVISGPIFDYDYGDLHDTEDKIKQYVEGSSIPVPTHYYSIITSCILDFTQPADKCDGP	840
QY	841	LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPE	900
Db	841	LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPE	900
QY	901	ILTLKTYLHTYSEI	915
Db	901	ILTLKTYLHTYSEI	915
RESULT 2			
A55453			
plasma cell membrane glycoprotein PC-1, brain specific - rat			
N;Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)			
C;Species: Rattus norvegicus (Norway rat)			
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000			
C;Accession: A55453; JU0187			
R;Narita, M.; Goji, J.; Nakamura, H.; Sano, K.			
J. Biol. Chem. 269, 28235-28242, 1994			
A;Title: Molecular cloning, expression, and localization of a brain-specific phosphodies			
A;Reference number: A55453; MUID:95050605			
A;Accession: A55453			
A;Molecule type: mRNA			
A;Residues: 1-885 <NAR>			
A;Cross-references: GB:D28560; NID:g464196; PIDN:BAA05910.1; PID:g464197			
R;Narita, M.; Goji, J.; Sano, K.; Nakamura, H.			
submitted to JIPID, February 1994			

A;Description: Cloning and expression of brain-specific nucleotide diphosphohydrolase  
A;Reference number: JU0187  
A;Accession: JU0187  
A;Molecule type: mRNA  
A;Residues: 1-66,'Q',68-81,'T',83-94,'C',96,'A',98-195,'A',197-514,'E',516-621,'E',62  
A;Experimental source: strain Sprague-Dawley  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: coenzyme A; exonuclease; glycoprotein; phosphoprotein; phosphoric diester  
F;54-97/Domain: somatomedin B homology <SBH1>  
F;98-141/Domain: somatomedin B homology <SBH2>  
F;53,150,396,408,522,608,829/Binding site: carbohydrate (Asn) (covalent) #status pred  
F;207/Binding site: AMP (Thr) (covalent) #status predicted

Query Match		85.8%;	Score 4306;	DB 1;	Length 885;
Best Local Similarity		83.6%;	pred. No. 1.7e-287;		
Matches 792;		Conservative 33;	Mismatches 28;	Indels 94;	Gaps 6;
QY	1	MARRSSFQSQIIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGSCK	60		
Db	1	MARQGCLGSFQVISLFTFAISVNICLGLFTASRIKRAE-WDEGPPTVLSDSPWTNTSGSK	59		
QY	61	GRCFELQEAQPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRGGEVRNEENACHC	120		
Db	60	GRCFELQEVGPPDCRCDNLCKSYSSCCHDFDELCLKTVRGWECTKDRSGEVRNEENACHC	119		
QY	121	SEDCLARGDCCTNYQVVKGESHVWDD-----DCEEIKAAECPAGFVRPPLIIFSVDG	173		
Db	120	PEDCLSRGDCCTNYQVVKGESHVWDDAARNQSSECLQV---CP-----PPLIIFSVDG	170		
QY	174	FRASYMKKGSKVMPIEKLKSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSM	233		
Db	171	FRASYMKKGSKVMPIEKLKSCGTHVPYTRPVYPTKTFPNLYTLATGLYPESHGIVGNSM	230		
QY	234	YDPVFDATFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPHERRLTILRWL	293		
Db	231	YDPVFDASFHLRGREKFNHRWGGQPLWITATKQGVRACTFFWSVIPHERRLTILQWL	290		
QY	294	TLPDHERPSVYAFYSEQPDFSGHKYGPFGPESSYSGSPFPAKRPKRKVAPKRQRPVA	353		
Db	291	SLPDNERPSVYAFYSEQPDFSGHKYGPFGPE-----	321		
QY	354	PPKRRRKIHRMDHYAAETRODKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHG	413		
Db	322	-----MTNPLREIDKTVGQLMDGLKQLRLHRCVNVIFVGDHG	358		
QY	414	MEDVTCDRTEFLSNLYLTNVDDITLVPGTLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHF	473		
Db	359	MEDVTCDRTEFLSNLYLTNVDDITLVPGTLGRIRAKSINNSKYDPKTIIANLTCKKPDQHF	418		
QY	474	KPYLKQHLPKRLHYANNRRIEDIHLVRRWHVARKPLDVYKKPSGKCFQDGHGFDNKV	533		
Db	419	KPYMKQHLPKRLHYANNRRIEDIHLVDRRWHVARKPLDVYKKPSGKCFQDGHGFDNKV	478		
QY	534	NSMQTVFVGYPGTFKYKTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRTNTFR	593		
Db	479	NSMQTVFVGYPGTFKYKTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRTNTFR	538		
QY	594	PTMPEEVRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDLNLKRLHTKSGTE-----	645		
Db	539	PTMPDEVSRPNYPGIMYLQSEFDLGCTCDDKVEPKNKEELNKLRLHTKSGTEAETGKFRG	598		
QY	646	-----ERHLLYGRPAVLRYTRYDILYHTDFESGYSEIFLMLLWTSYTV	688		
Db	599	SKHENKKNLNGSVEPRKERHLLYGRPAVLRYTSYDILYHTDFESGYSEIFLMLPLWTSYTI	658		
QY	689	SKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCLAYKNDKQMSYGFLFPPYLSSSSPEAKYDA	748		
Db	659	SKQAEVSSIPEHLTNCVRPDVVRVSPGFSQNCLAYKNDKQMSYGFLFPPYLSSSSPEAKYDA	718		
QY	749	FLVTNMVPMYPAFKRVWNVFQRLVKKYASERNGVNVISGPIFDYDYGDLHDTEDKIKQY	808		
Db	719	FLVTNMVPMYPAFKRVWAYFQRLVKKYASERNGVNVISGPIFDYDGLRDTDEDEIKQY	778		

QY 809 VEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWV 868  
Db 779 VEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWV 838  
QY 869 EELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915  
Db 839 EELMKMHTARVRDIEHLTGLDFYRKTSRSYSEILTLKTYLHTYESEI 885  
RESULT 3  
A57080  
cell surface antigen RB13-6 - rat  
N:Contains: phosphodiesterase I (EC 3.1.4.1)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A57080  
R:Deissler, H.; Lottspeich, F.; Rajewsky, M.F.  
J. Biol. Chem. 270, 9849-9855, 1995  
A:Title: Affinity purification and cDNA cloning of rat neural differentiation and tumor  
A:Reference number: A57080; MUID:95247775  
A:Accession: A57080  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-875 <DEI>  
A:Cross-references: GB:247987; NID:g806378; PIDN:CAA88029.1; PID:g806379  
C:Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C:Keywords: exonuclease; glycoprotein; phosphoprotein; phosphoric diester hydrolase; su  
F:1-22/Domain: cytosolic #status predicted <CYT>  
F:23-45/Domain: transmembrane #status predicted <TMM>  
F:46-875/Domain: extracellular #status predicted <EXT>  
F:51-94/Domain: somatomedin B homology <SBH>  
F:95-138/Domain: somatomedin B homology <SBH2>  
F:206/Binding site: AMP (Thr) (covalent) #status predicted  
F:237,280,289,533,574,594,702,789/Binding site: carbohydrate (Asn) (covalent) #status pr  
Query Match 41.5%; Score 2084; DB 1; Length 875;  
Best Local Similarity 43.5%; Pred. No. 5e-135;  
Matches 386; Conservative 160; Mismatches 248; Indels 94; Gaps 11;  
QY 57 GSCKGRCFELQEGPPDCRDNLCKSYTSCCHDFDELCKLTARGWECTKDRCGEVRNEEN 116  
Db 53 GSCRKKCFDSSHRGLEGCRCDSGCTDRGCCWDFEDTCVKSTQIWTCSNFRCGETRLEAA 112  
QY 117 ACHCEDCLARGDCCTNYQVCKGESHVDDDCDEEIKAAECAPGFVRPPLIIFSVDFERA 176  
Db 113 LCSCADDCLQRKDCCTDYKAVCQGEVPWVTEACASSQEPQCEGFDQPPVILFSMDGFERA 172  
QY 177 SYMKKGSKVPNIEKLRSCTGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDP 236  
Db 173 EYLQWSTLLPNINKLTKCGLHSKYMAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDV 232  
QY 237 VFDAFHLRGREFNHRWGGQPLWITATKQGVKAGTFFW-----SV 278  
Db 233 YLNKNFSLSSVEKSNPAWWSGQPIWLTAMYQGLKAASYWLGSVDVAVNGSFNPNIYRYSN 292  
QY 279 VIPHERRILTLRWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPESSYGSFPFPAKRP 338  
Db 293 SVPYESRIATLLQWLDLPKAERPSFYTIYVEEPSAGHSGPV-----SAG----- 338  
QY 339 KRKVAPKRRQERPVPAPPKRRRKIHRMDHYAAETRQDKMTNPLREIDKIVGQMDGLKQL 398  
Db 339 -----VIKALQLVDDAFGLMEGLKOR 360  
QY 399 KLRRCVNVIFVGDHGMEDVTCDRTEFLSNLYTNVDDITLVPGLGRIRSK-----FSNN 452  
Db 361 NLHNCVNIIVLADHGMQDTSRDRVEYMTDYFPEI-NFYMYQGPAPRIRTNIPQDFFTN 419  
QY 453 AKYDPKAIIANLTCKKPDQHFQKPYLKQHLPKRLHYANNRRIEDIHLVVERRWHVARKPLD 512  
Db 420 S----EEIVRDLSCRKSDQHFKPYLTPDLPKRLHYAKNVRIDKVLMLVDRQWLAYR---- 471

QY 513 VYKKPSGKCFQGDHGFNDKNVSMQTVFVGYGPTFKYKTKVPPFENIELYNVMDLLGLK 572  
Db 472 --NKGSSNC-EGGTHGYNNEFKSMEAIFLAHGSPSKKTKVIEPFENIEVYNLDCDLHLIQ 528  
QY 573 PAPNNGTHGSLNHLRTNTRPTMPEEVRPNYPNGIMYLSQDFDLGCTCDDKVPEKNKLD 632  
Db 529 PAPNNGSHGSLNHLKAPFYQPSHAEELSKSAGCGFTPLPKDSLNCSC-LALQTSQEE 587  
QY 633 ELNKRLLHTKG----STEERHLLYGRPAVLYRTR-YDILYHTDFESGYSEIFLMLLWTSYT 687  
Db 588 QVNQRLNLSGGEVSATEKTNLPFGRPRVIOKNKHCLLYHREYVSGFGKAMKMPMWSSYT 647  
QY 688 VSKQAEVSSVPDHLTSCVRPDVRVSPFSQNCLAYKNDKQMSYGFLFPPPYLSSSPKAYD 747  
Db 648 VPKPGDTSSLPTVPDCLRADVRVDPSESKCSFYLADQNDIDHGFLYPPIAIGNNESQYD 707  
QY 748 AFLVTNMVPMYPAPKRVWNYFORVLVKKYASERNGVNVISGPIFDYDYGDLHDTEDKIKQ 807  
Db 708 ALITSNLVPMYKEFKMWDYFHKVLLIKYAIERNGVNVVSGPIFDYNYDGHFADPEITN 767  
QY 808 YVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKW 867  
Db 768 YVAGTDVPVPTHYFVLTSCNKTHTPDSCPGWLDVLPFVVPHRPTNVESCPENKAEDLW 827  
QY 868 VEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915  
Db 828 VEERFKAHARVRDVELLTGLDFYQEKTPQVSEILQLKTYLPTFETII 875  
RESULT 4  
A39216  
plasma cell membrane glycoprotein PC-1 - human  
N:Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-May-2000  
C:Accession: A39216; S21706; S23587; S51030  
R:Buckley, M.F.; Loveland, K.A.; McKinstry, W.J.; Garson, O.M.; Goding, J.W.  
J. Biol. Chem. 265, 17506-17511, 1990  
A:Title: Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human molecule,  
A:Reference number: A39216; MUID:91009202  
A:Accession: A39216  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-925 <BUC>  
A:Cross-references: GB:J05654  
R:Funakoshi, I.; Kato, H.; Horie, K.; Yano, T.; Hori, Y.; Kobayashi, H.; Inoue, T.; S  
Arch. Biochem. Biophys. 295, 180-187, 1992  
A:Title: Molecular cloning of cDNAs for human fibroblast nucleotide pyrophosphatase.  
A:Reference number: S21706; MUID:92246539  
A:Accession: S21706  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-925 <FUN1>  
A:Accession: S23587  
A:Molecule type: protein  
A:Residues: 116-121;247-271,'X',273-275;279-280,'X',282-283;303-316;362-364;449-465;4  
A>Note: it is uncertain whether Met-1 or Met-53 is the initiator  
R:Belli, S.I.; Goding, J.W.  
Eur. J. Biochem. 226, 433-443, 1994  
A:Title: Biochemical characterization of human PC-1, an enzyme possessing alkaline ph  
A:Reference number: S51030; MUID:95094801  
A:Accession: S51030  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-80 <BEL>  
C:Genetics:  
A:Gene: GDB:PDNPL1; M6S1; NPPS  
A:Cross-references: GDB:132615; OMIM:173335  
A:Map position: 6q22-6q23  
C:Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C:Keywords: coenzyme A; glycoprotein; phosphoprotein; phosphoric diester hydrolase; t  
F:77-97/Domain: transmembrane #status predicted <TMM>





QY 458 KAIIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRRIEDIHLLVERRWHVARKPLDVYKPK 517  
Db 454 EALAKNLSCREPNHFRPYLKPFLPKRLHFAKSDRIEPLTFYLDPOWQLALNPSE--RKY 511  
QY 518 SGKCFQGDHGFNDKVNVMQTVFVGYGPTFKYKTKVPPFENIELYNVCMCDLLGLKPPAPNN 577  
Db 512 CG----SGFHGSDNLFNSMQALFIGYGPAFKHGAENVDSFENIEVYNLMCDLLGLIPAPNN 567  
QY 578 GTHGSLNHLRTNFRPTMPPEEVRPNYPGIM---YLQSDFDLGCTCDDKVEPKNKLD 633  
Db 568 GSHGSLNHLKKPIYNPSHPKE-----EGFLSQCPKIKSTSDNLGCTCDPWIVP---IKD 618  
QY 634 LNKRLHTKGSTEERHLL-----YGRPAVLYRTRY-DILYHTDFESGYSEIFLMLLWTSYT 687  
Db 619 FEQQLNL--TTEDDDIYHMTVPYGGPRILLKQHVVCLLQOQQFLTGYSLDLLMPLWASYT 676  
QY 688 VSKQAEVSSVPDHLTSCVRPDVRVSPFSQNCCLAYKNDKQMSYGFLFPPYLSSSPEAKY- 746  
Db 677 FLRNDQFSR--DDFSCNMQDLRIPLSPVHKCSYKSKLSYGFLTTPPRLNRVSNHIYS 734  
QY 747 DAFLVTNMVMPYPAFKRVWNYFORVLVKKYASERNGVNVISGPIDFYDYDGLHDTEDKIK 806  
Db 735 EALLTSNIVPMYQSFQVIWHYHLHDTLLQRYAHERNGINVVSGPVDFDYDGRYDSLEILK 794  
QY 807 Q---YVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSED 863  
Db 795 QNSRVIRSQEILIPTHFFIVLTCKQLSETPLECSA-LESSAYILPHRPDNIESCETHGKR 853  
QY 864 ESKWVEELMKHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESE 914  
Db 854 ESSWVEELLTLHRAVRTDVELITGLSFYQDRQESVSELLRLKTHLPFSQE 904  
RESULT 6  
T09932  
nucleotide pyrophosphatase homolog T16L4.200 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999  
C:Accession: T09932  
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, June 1999  
A:Reference number: Z16897  
A:Accession: T09932  
A:Molecule type: DNA  
A:Residues: 1-457 <BEV>  
A:Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.200  
A:Experimental source: cultivar Columbia; BAC clone T16L4  
C:Genetics:  
A:Gene: ATSP:T16L4.200  
A:Map position: 4  
Query Match 12.9%; Score 646.5; DB 2; Length 457;  
Best Local Similarity 32.0%; Pred. No. 9.3e-37;  
Matches 151; Conservative 78; Mismatches 136; Indels 107; Gaps 12;  
QY 146 DDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKSGKVMPIEKLRSCTHSPY-MRP 204  
Db 37 DSPSSYVRRPQPKLKNKPVLLISCDGFRFGYQFTE--TPNIDLLISRGTEAKTGLIP 94  
QY 205 VYPTKTFPNLYLATGLYPESHGIVGNSMYDPVDFATFHLRGREKENHRWVGQQLWITA 264  
Db 95 VFPTMTFPNHYSIATGLYPASHGIIMNKFTDPVSGELFN----RNLNPKWWLGEPLWTA 150  
QY 265 TKQGVKAGTFW-----SVVIPHERILLTILRWLTLPDHERPSV 303  
Db 151 VNQGLMAATYFWPGADVHKGSWNCPKGFCAPYNSVPLEERVDTILNYFDLPERETPDF 210  
QY 304 YAFYSEQPDFSGHKYGPFGPESSYGSPTPAKRPKRKVAPKRRQERPVPAPKKRRRKIH 363  
Db 211 MALYFDEPDIOGHEYGPDDP----- 230

QY 364 RMDHYAAETRODKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTE 423  
Db 231 -----RVTEAVSKVDKMGIRIIMGLEKRKRVFSDVHVILLGDHGMV-TNCDKKV 277  
QY 424 FLSNYLTNVDDITLVPG-----TLGR-IRSKFSNNAKYDPKAIIANLTCKK 468  
Db 278 I--YIDDLADWIKIPADWIDYSPVLAMNPRWGDKVKNPGQKNAELVRKMNEALSSGKV 334  
QY 469 PDQHF-KPYLKQHLPKRLHYANNRRRIEDIHLLVERWHVARKPLDVYKKPSGKCFQGDH 527  
Db 335 ANGEFLQVYLKENLPQRLHYSDSSRIPIIGMVGEGLMVKQNRITYQ-----EC--SGTH 387  
QY 528 GFDNKVNSMQTVFVGYGPTFKYKTKVPPFENIELYNVCMCDLLGLKPPAPNNGT 579  
Db 388 GYDNMFFSMRSIFVGYGPRFRRGIKVPSFENVQYNAVAEILGLRPAPNNGS 439  
RESULT 7  
T09931  
nucleotide pyrophosphatase homolog T16L4.190 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999  
C:Accession: T09931  
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.  
submitted to the Protein Sequence Database, June 1999  
A:Reference number: Z16897  
A:Accession: T09931  
A:Molecule type: DNA  
A:Residues: 1-496 <BEV>  
A:Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.190  
A:Experimental source: cultivar Columbia; BAC clone T16L4  
C:Genetics:  
A:Gene: ATSP:T16L4.190  
A:Map position: 4  
Query Match 12.2%; Score 612; DB 2; Length 496;  
Best Local Similarity 32.7%; Pred. No. 2.4e-34;  
Matches 147; Conservative 68; Mismatches 131; Indels 104; Gaps 12;  
QY 163 RPPLIIFSVDGFRASYMKKSGKVMPIEKLRSCTHSPY-MRPVYPTKTFPNLYTLATGL 221  
Db 106 KPVVLLISSDGFREFGYQFTKK--LPSIHRLLIANGTEAETGLIPVFTLTFPNHYSIVTGL 163  
QY 222 YPESHGIVGNSMYDPVDFATFHLRGREKENHRWVGQQLWITATKQGVKAGTFW----- 276  
Db 164 YPAYHGIINNHFVDPETGNVFTMASHEP---EWWLGEPLWETVNVNQGKAAATYFWPGSEV 220  
QY 277 -----SVVIPHERILLTILRWLTLPDHERPSVYAFYSEQPDFSGHKYGP 321  
Db 221 HKGSWNCPOGLCQNYGVSVPFDDRVDTILSYFDLPNEIPSFMTLYFEDPDHGHQVGP 280  
QY 322 GPESSYGSPTPAKRPKRKVAPKRRQERPVPAPKKRRRKIHRMDHYAAETRODKMTNPL 381  
Db 281 DPQ-----ITEAV 288  
QY 382 REIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNLYTNVDDIT---LV 438  
Db 289 VNIDRLIGRLIDGLEKRGIFEDVTMIMVGDHGMVG-TCDKK-----LVVLDLAPWIKI 341  
QY 439 PGT-----LGRIRSKFSNNAKYDPKAIIANLTCKKPD--QHEKPYLKQHLPKRLHYAN 489  
Db 342 PSSWVQYTYTLLATQPPSGHDAADIVAKINEGLSSGKVENGKYLKVYLKEDLP SRLHYVD 401  
QY 490 NRRIEDIHLLVERWHVARKPLDVYKKPSGKCFQGDHGFNDKVNMSQTVFVGYGPTFKY 549  
Db 402 SDRIPIIGLVDEGFKVEQK-----KSKAKEC--GGAHGYDNAFFSMRTIFIGHGPMFSK 454  
QY 550 KTKVPPFENIELYNVCMCDLLGLKPPAPNNGT 579  
Db 455 GRKVPSEFENVQIYNVISSILGLKAAAPNNGS 484









QY 665 LYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVR-----VSPSFSQN 718  
Db 518 --NNCFQS-YCE-----NSLIHKN-----RQDVRKGVIESLTFFSRN 553  
QY 719 CLAYKNDQKMSYGFLEPPPY-----LSSSPEAKY---DAFLVTMV 755  
Db 554 QSVFEN---SFSFVNTKYSIECPKLDTKDNFFTAGSEAIKLANAQYKFPSPFMKSELI 609  
QY 756 PMYPAFK-----RVWNYFQORVLVKKYASERNGVNVISGPIFDYDGLHDTED 803  
Db 610 SSLLSLKDETIKFVDIWPVLS-IKTDEYLKHYGKLFVLSGLAVDRNLDGIADDEE 663

RESULT 15  
A41179  
protein kinase PC-1 (EC 2.7.1.1) - bovine (fragments)  
N:Alternate names: MAFP; major acidic fibroblast growth factor-stimulated phosphoprotein  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 28-May-1992 #sequence\_revision 22-Apr-1995 #text\_change 30-Apr-1999  
C:Accession: A41179; A49308  
R:Oda, Y.; Kuo, M.D.; Huang, S.S.; Huang, J.S.  
J. Biol. Chem. 266, 16791-16795, 1991  
A:Title: The plasma cell membrane glycoprotein, PC-1, is a threonine-specific protein kinase  
A:Reference number: A41179; MUID:91358477  
A:Accession: A41179  
A:Molecule type: protein  
A:Residues: 1-26;36-56;59-67;68-133;134-144;145-267;268-300 <ODA>  
A:Experimental source: liver  
R:Oda, Y.; Kuo, M.D.; Huang, S.S.; Huang, J.S.  
J. Biol. Chem. 268, 27318-27326, 1993  
A:Title: The major acidic fibroblast growth factor (aFGF)-stimulated phosphoprotein from  
estrase activities.  
A:Reference number: A49308; MUID:94086550  
A:Accession: A49308  
A:Molecule type: protein  
A:Residues: 27-35,'X',37-58 <OD2>  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBIP:141583)  
C:Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C:Keywords: glycoprotein; phosphoprotein; phosphotransferase  
F:1-25/Domain: somatomedin B homology (fragment) <SBH>  
F:36/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 9.1%; Score 456; DB 2; Length 300;  
Best Local Similarity 18.7%; Pred. No. 6.3e-24;  
Matches 150; Conservative 45; Mismatches 86; Indels 522; Gaps 14;

QY 107 RCGEVRNEENACHCEDCLARGDCCTNYQVCKGESHVDDDCCEEIKAAECPAGFVRPPL 166  
Db 2 RCFE-RTFGN-CRCDAACVELGNCCLD----- 26

QY 167 IIFSVDGFRASYNMKGSKVMPNIEKLRCGTHSPYMRPVYPTKTFPNLYTLATGLYPESH 226  
Db 27 -----NMRPVYPTKTFPNHYSIVTGLYPESH 52

QY 227 GIVGNSMYDPVDFATFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFW----- 276  
Db 53 GIIDNK-----GEPWLTA-----KSGTFFWPGSDVKINGI 83

QY 277 -----SVVIPHERILTILRWLTLPDHERPSVYAFYSEQPDFSGHKYGFPGPESSYGSP 331  
Db 84 FPDYXVSVPFEEILAILKWQLP----- 108

QY 332 FTPAKRPRKRVAPKRRQERPAPPKRRRKIHRMDHYAAETRODKMTNPLREIDKIVGQL 391  
Db 109 ----- 108

QY 392 MDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNLYTNVDDITLVPGTIGRIRSKFSN 451  
Db 109 ----KELNLHRCNLILISDHGMEQGSCKYYSF-----DYEG 141

QY 452 NAKYDPKAIITANLTCKKPDQHFKPYLKQHLPKRLHYANNRRRIEDIHLLVERHWHVARKPL 511  
Db 142 IAKSD-----RIERLTFYLDPOWQALNPS 166

QY 512 DVYKKPSGKCFQGDHGFNDKVNSMQTVFVGYGPTFKYKTKVPPFENIELNVNMCDDLGL 571  
Db 167 E--RKYCG---GGFHGSDNLFNMQALFXXXXXXHSTEVDSFENIEVNLNMCDDLNL 220

QY 572 KPAPNNGTHGSLNHLRLTNTFRPTMPEEVTRPNYPGIMYLQSDFDLGCTCDDKVEPKNKL 631  
Db 221 TPAPNNGTH----- 229

QY 632 DELNKRLLHTKGSTEEHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTVSKQ 691  
Db 230 ----- 229

QY 692 AEVSSVPDHLTSCVRPDVRVSPFSQNCLAYKNDKQMSYGFLEPPYLSSSPEAKY-DAFL 750  
Db 230 -----LSYGFLSPQLHKGSSQVYSEALL 253

QY 751 VTNMVPMYPAPKRVWNYFORVLVKKYASERNGVNVISGPIFDYDGLHDTEDKIKQYVE 810  
Db 254 TTNIVPMYQSFOVI----- 267

QY 811 GSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEE 870  
Db 268 -----HESLWVEE 275

QY 871 LMKMHTARVRDIEHLSLDFFRK 893  
Db 276 LLKLHTARITDVEHITGLSFYQQ 298

Search completed: July 19, 2001, 14:45:54  
Job time: 152 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2001, 14:44:18 ; Search time 16.91 Seconds  
(without alignments)  
1853.564 Million cell updates/sec

Title: US-09-483-831-69  
Perfect score: 5019  
Sequence: 1 MARRSSFQSQIISLTFV.....RSTYPEILTKYLHTESEI 915

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1955.5	39.0	873	1	PC1_HUMAN	P22413 homo sapien
2	1899.5	37.8	871	1	PC1_MOUSE	P06802 mus musculu
3	410	8.2	743	1	YCR6_YEAST	P25353 saccharomyc
4	335.5	6.7	493	1	YEB6_YEAST	P39997 saccharomyc
5	238	4.7	61	1	PPD1_BOVIN	P15396 bos taurus
6	146.5	2.9	329	1	NUC1_YEAST	P08466 saccharomyc
7	123	2.5	4655	1	LRP2_HUMAN	P98164 homo sapien
8	122	2.4	1323	1	LT23_CAEEL	P24348 caenorhabdi
9	121	2.4	1888	1	YDT2_SCHPO	O14207 schizosacch
10	119.5	2.4	252	1	NUC1_CUNEE	P81203 cunninghame
11	119.5	2.4	475	1	VTNC_RABIT	P22458 oryctolagus
12	117	2.3	335	1	NUC1_SCHPO	Q10480 schizosacch
13	116.5	2.3	989	1	YD30_YEAST	P32898 saccharomyc
14	112	2.2	1394	1	TGFB_HUMAN	P22064 homo sapien
15	111.5	2.2	299	1	NUCG_BOVIN	P38447 bos taurus
16	111	2.2	1396	1	ITA2_DROME	P12080 drosophila
17	111	2.2	2871	1	FBN1_HUMAN	P35555 homo sapien
18	110.5	2.2	294	1	NUCG_MOUSE	O08600 mus musculu
19	110.5	2.2	478	1	VTNC_MOUSE	P29788 mus musculu
20	110.5	2.2	2871	1	FBN1_BOVIN	P98133 bos taurus
21	110	2.2	753	1	YJ05_CAEEL	O01975 caenorhabdi
22	109.5	2.2	1472	1	ATC9_YEAST	Q12697 saccharomyc
23	108.5	2.2	320	1	NUC1_SYNRA	P81204 syncephalas
24	108	2.2	1065	1	YD22_SCHPO	O13710 schizosacch
25	106	2.1	1106	1	STC_DROME	P40798 drosophila
26	106	2.1	1178	1	YNI7_YEAST	P48231 saccharomyc
27	106	2.1	2204	1	RRPL_NDV8	P11205 newcastl d
28	106	2.1	2871	1	FBN1_MOUSE	Q61554 mus musculu
29	106	2.1	3099	1	POLG_PEMVM	O56075 p genome po
30	105.5	2.1	1612	1	MTDM_PARLI	Q27746 paracentro
31	105.5	2.1	1808	1	TENA_CHICK	P10039 gallus gall
32	104.5	2.1	2201	1	TENA_HUMAN	P24821 homo sapien
33	104.5	2.1	4289	1	TENX_HUMAN	P22105 homo sapien

34	104	2.1	470	1	BFR1_YEAST	P38934 saccharomyc
35	104	2.1	659	1	AMIA_STRPN	P18791 streptococc
36	104	2.1	700	1	KPC2_DROME	P13677 drosophila
37	104	2.1	969	1	DPOM_NEUIN	P33538 neurospora
38	103.5	2.1	297	1	NUCG_HUMAN	Q14249 homo sapien
39	103	2.1	755	1	P100_HSV7J	P52519 human herpe
40	103	2.1	1700	1	BAR3_CHITE	Q03376 chironomus
41	102.5	2.0	369	1	PH11_HUMAN	P21128 homo sapien
42	102.5	2.0	747	1	FHUA_ECOLI	P06971 escherichia
43	102.5	2.0	1426	1	EGFR_DROME	P04412 drosophila
44	102	2.0	563	1	IDS_MOUSE	Q08890 mus musculu
45	102	2.0	1018	1	FNBA_STAAU	P14738 staphylococ

ALIGNMENTS

RESULT 1						
PC1_HUMAN		ID	PC1_HUMAN	STANDARD;	PRT;	873 AA.
		AC	P22413; Q9Y6K3; Q9UP61;			
		DT	01-AUG-1991 (Rel. 19, Created)			
		DT	01-AUG-1991 (Rel. 19, Last sequence update)			
		DT	30-MAY-2000 (Rel. 39, Last annotation update)			
		DE	PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 [INCLUDES: ALKALINE			
		DE	PHOSPHODIESTERASE I (EC 3.1.4.1); NUCLEOTIDE PYROPHOSPHATASE			
		DE	(EC 3.6.1.9) (NPPASE)].			
		GN	PDNP1 OR PC1 OR NPPS.			
		OS	Homo sapiens (Human).			
		OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
		OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
		OX	NCBI_TaxID=9606;			
		RN	[1]			
		RP	SEQUENCE FROM N.A.			
		RX	MEDLINE=91009202; PubMed=2211644;			
		RA	Buckley M.F., Loveland K.A., McKinstry W.J., Garson O.M., Goding J.W.;			
		RT	"Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human			
		RT	molecule, amino acid sequence, and chromosomal location.";			
		RL	J. Biol. Chem. 265:17506-17511(1990).			
		RN	[2]			
		RP	SEQUENCE FROM N.A.			
		RX	MEDLINE=92246539; PubMed=1315502;			
		RA	Funakoshi I., Kato H., Horie K., Yano T., Hori Y., Kobayashi H.,			
		RA	Inoue T., Suzuki H., Fukui S., Tsukahara M., Kajii T.,			
		RA	Yamashina I.;			
		RT	"Molecular cloning of cDNAs for human fibroblast nucleotide			
		RT	pyrophosphatase.";			
		RL	Arch. Biochem. Biophys. 295:180-187(1992).			
		RN	[3]			
		RP	SEQUENCE OF 93-133 FROM N.A., AND VARIANT GLN-121.			
		RX	MEDLINE=99408501; PubMed=10480624;			
		RA	Pizzuti A., Frittitta L., Argiolas A., Baratta R., Goldfine I.D.,			
		RA	Bozzali M., Ercolino T., Scarlato G., Iacoviello L., Vigneri R.,			
		RA	Tassi V., Trischitta V.;			
		RT	"A polymorphism (K121Q) of the human glycoprotein PC-1 gene coding			
		RT	region is strongly associated with insulin resistance.";			
		RL	Diabetes 48:1881-1884(1999).			
		CC	-!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.			
		CC	-!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES			
		CC	SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED			
		CC	OLIGO-NUCLEOTIDES.			
		CC	-!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.			
		CC	-!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.			
		CC	-!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.			
		CC	-!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER			
		CC	OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE			
		CC	OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.			
		CC	-!- SIMILARITY: CONTAINS 2 SOMATOMEDIN-B TYPE DOMAINS.			
		CC	-----			
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
		CC	the European Bioinformatics Institute. There are no restrictions on its			
		CC	use by non-profit institutions as long as its content is in no way			











RP SEQUENCE.  
RC TISSUE=Intestine;  
RX MEDLINE=85234541; PubMed=2989287;  
RA Culp J.S., Blytt H.J., Hermodson M., Butler L.G.;  
RT "Amino acid sequence of the active site peptide of bovine intestinal  
RT 5'-nucleotide phosphodiesterase and identification of the active  
RT site residue as threonine.";  
RL J. Biol. Chem. 260:8320-8324(1985).  
CC -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES  
CC SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED  
CC OLIGO-NUCLEOTIDES.  
DR PIR; A25274; A25274.  
KW Hydrolase.  
FT NON\_TER 1 1  
FT ACT\_SITE 39 39 FORMS A PHOSPHOTHEONINE INTERMEDIATE.  
SQ SEQUENCE 61 AA; 6999 MW; 580C8B4807A61C84 CRC64;

Query Match 4.7%; Score 238; DB 1; Length 61;  
Best Local Similarity 66.7%; Pred. No. 5.4e-11;  
Matches 40; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 172 DGFRAASYMKGSKVMPNIEKLRSCGTHSPYMRPVYPTKTPNLYTLATGLYPESHGIVGN 231  
Db 1 DGFRAEYLQWTSTLVPNINKLTCGVHSQYLRPAYPTKTPPNHYTIVTGLYPESHGIIDN 60

RESULT 6  
NUCL\_YEAST  
ID NUCL\_YEAST STANDARD; PRT; 329 AA.  
AC P08466;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE MITOCHONDRIAL NUCLEASE (EC 3.1.30.-).  
GN NUC1 OR YJL208C OR J0310 OR HRE329.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=2935-106;  
RX MEDLINE=88233924; PubMed=2836792;  
RA Vincent R.D., Hofmann T.J., Zassenhaus H.P.;  
RT "Sequence and expression of NUC1, the gene encoding the mitochondrial  
RT nuclease in Saccharomyces cerevisiae.";  
RL Nucleic Acids Res. 16:3297-3312(1988).  
RN [2]  
SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=95274326; PubMed=7754713;  
RA Purnelle B., Coster F., Goffeau A.;  
RT "The sequence of a 36 kb segment on the left arm of yeast chromosome  
RT X identifies 24 open reading frames including NUC1, PRP21 (SPP91),  
RT CDC6, CRY2, the gene for S24, a homologue to the aconitase gene ACO1  
RT and two homologues to chromosome III genes.";  
RL Yeast 10:1235-1249(1994).  
RN [3]  
SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=95242842; PubMed=7725802;  
RA Vandenbol M., Durand P., Bolle P.-A., Dion C., Portetelle D.,  
RA Hilger F.;  
RT "Sequence analysis of a 40.2 kb DNA fragment located near the left  
RT telomere of yeast chromosome X.";  
RL Yeast 10:1657-1662(1994).  
RN [4]  
SEQUENCE OF 204-329 FROM N.A.  
RX MEDLINE=90014786; PubMed=2552292;  
RA Liu Y., Dieckmann C.L.;  
RT "Overproduction of yeast viruslike particles by strains deficient in  
RT a mitochondrial nuclease.";

RL Mol. Cell. Biol. 9:3323-3331(1989).  
CC -!- FUNCTION: THIS ENZYME HAS BOTH RNASE AND DNASE ACTIVITY.  
CC -!- COFACTOR: REQUIRES MANGANESE OR MAGNESIUM.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE DNA/RNA NON-SPECIFIC ENDONUCLEASES  
CC FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; X06670; CAA29870.1; -.  
DR EMBL; X77688; CAA54748.1; -.  
DR EMBL; Z34098; CAA84003.1; -.  
DR EMBL; Z49483; CAA89505.1; -.  
DR EMBL; M28067; AAA34457.1; -.  
DR PIR; S05888; NCBYN1.  
DR PIR; S46621; S46621.  
DR HSSP; P13717; 1SMN.  
DR SGD; S0003744; NUC1.  
DR InterPro; IPR001604; -.  
DR Pfam; PF01223; Endonuclease; 1.  
DR PROSITE; PS01070; NUCLEASE\_NON\_SPEC; 1.  
KW Hydrolase; Nuclease; Endonuclease; Magnesium; Mitochondrion;  
KW Inner membrane.  
FT ACT\_SITE 138 138 BY SIMILARITY.  
SQ SEQUENCE 329 AA; 37209 MW; CAEC9678CB00943C CRC64;

Query Match 2.9%; Score 146.5; DB 1; Length 329;  
Best Local Similarity 22.8%; Pred. No. 0.002;  
Matches 79; Conservative 48; Mismatches 159; Indels 61; Gaps 16;

QY 565 MCDLLGLKAPANNNGTHGSLNHLRTNTFRPTMPEEVTRPNYPGIMYLSQDFDLGCTCDDK 624  
Db 1 MCSRILLGLVGLGAGTGLTYLL-LNKHSTP---QIETPYP----- 38

QY 625 VEPKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLWT 684  
Db 39 --PTQKPNS-NIQSHSFNVDPSPGFFKYGFGPI----HDLQNRREEFISCYNRQTQNPYVW 91

QY 685 SYTVSKQAEVSSVPDHLTSCVRPDRVSPFSQNCCLAYKNDKQMSYGFLEPPY-LSSSPE 743  
Db 92 LEHITPESLAARNADRNKNSFFKEDEVIPEKFRGLRDYFR-SGYDRGHQAPAAADAKFSQQ 150

QY 744 AKYDAFLVTNMVPMY-PAFKR-VW---NYFQRLVKKYASERNGVNIVSGPIFDYDGL 798  
Db 151 AMDDTFYLSNMCPOVGEFNRDYNWAHLEYFCRGLTKKYKS---VRIVTGPLYLPKKDPI 206

QY 799 HDTEDKIKQYVEGS--SIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRP-DNE 855  
Db 207 -DNKFRVNYEVIGNPPSIAVPTHFFKLIVAEAPTANPARE---DIAVAAFVLPNEPISNE 262

QY 856 ESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSSYPEIL 902  
Db 263 TKLTDFE-----VPIDALERSTGLELLQKVPPSKKAL 295

RESULT 7  
LRP2\_HUMAN  
ID LRP2\_HUMAN STANDARD; PRT; 4655 AA.  
AC P98164; O00711; Q16215;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 2 PRECURSOR (MEGALIN)  
DE (GLYCOPROTEIN 330).  
GN LRP2.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=96305376; PubMed=8706697;  
RA Hjaelm G., Murray E., Crumley G., Harazim W., Lundgren S., Onyango I.,  
RA Ek B., Larsson M., Juhlin C., Hellman P., Davis H., Akerstroem G.,  
RA Rask L., Morse B.;  
RT "Cloning and sequencing of human gp330, a Ca(2+)-binding receptor  
RT with potential intracellular signaling properties.";  
RL Eur. J. Biochem. 239:132-137(1996).  
RN [2]  
RP SEQUENCE OF 2705-4453 FROM N.A.  
RC TISSUE=Kidney;  
RA Knaak C., Argraves W.S.;  
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 3833-4453 FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=95048397; PubMed=7959795;  
RA Korenberg J.R., Argraves K.M., Chen X.N., Tran H.,  
RA Strickland D.K., Argraves W.S.;  
RT "Chromosomal localization of human genes for the LDL receptor family  
RT member glycoprotein 330 (LRP2) and its associated protein RAP  
RT (LRPAP1).";  
RL Genomics 22:88-93(1994).  
RN [4]  
RP SEQUENCE OF 4139-4406 FROM N.A.  
RX MEDLINE=94244704; PubMed=8187828;  
RA Lundgren S., Hjaelm G., Hellman P., Ek B., Juhlin C., Rastad J.,  
RA Klarekog L., Akerstrom G., Rask L.;  
RT "A protein involved in calcium sensing of the human parathyroid and  
RT placental cytotrophoblast cells belongs to the LDL-receptor protein  
RT superfamily.";  
RL Exp. Cell Res. 212:344-350(1994).  
RN [5]  
RP FUNCTION.  
RX MEDLINE=95286588; PubMed=7768901;  
RA Kounnas M.Z., Loukinova E.B., Stefansson S., Harmony J.A.K.,  
RA Brewer B.H., Strickland D.K., Argraves W.S.;  
RT "Identification of glycoprotein 330 as an endocytic receptor for  
RT apolipoprotein J/clusterin.";  
RL J. Biol. Chem. 270:13070-13075(1995).  
CC -!- FUNCTION: BINDS SPECIFICALLY CLUSTERIN WITH HIGH AFFINITY, BUT  
CC ALSO LIGANDS IN COMMON WITH OTHER FAMILY MEMBERS: PLASMINOGEN,  
CC EXTRACELLULAR MATRIX COMPONENTS, PLASMINOGEN ACTIVATOR-PLASMINOGEN  
CC ACTIVATOR INHIBITOR TYPE I COMPLEX, APOLIPOPROTEIN E-ENRICHED  
CC BETA-VLDL, LIPOPROTEIN LIPASE, LACTOFERRIN AND CALCIUM.  
CC -!- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS  
CC APROTININ, AMINOGLYCOSIDES AND POLYMYXIN B (BY SIMILARITY).  
CC -!- FUNCTION: MAY PARTICIPATE IN REGULATION OF PARATHYROID-HORMONE AND  
CC PARA-THYROID-HORMONE-RELATED PROTEIN RELEASE.  
CC -!- SUBUNIT: FORMS A MULTIMERIC COMPLEX TOGETHER WITH A RECEPTOR-  
CC ASSOCIATED PROTEIN (RAP).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- TISSUE SPECIFICITY: ABSORPTIVE EPITHELIA, INCLUDING RENAL  
CC PROXIMAL TUBULES.  
CC -!- SIMILARITY: CONTAINS 36 LDL-RECEPTOR CLASS A DOMAINS.  
CC -!- SIMILARITY: CONTAINS 37 LDL-RECEPTOR CLASS B DOMAINS.  
CC -!- SIMILARITY: CONTAINS 17 EGF-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 4 SH3-BINDING DOMAINS.  
CC -!- SIMILARITY: CONTAINS 1 SH2-BINDING DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR EMBL; U33837; AAB41649.1; -  
DR EMBL; U04441; AAB02882.1; -  
DR EMBL; S73145; AAB30825.1; -  
DR HSSP; P01130; IAJJ.  
DR MIM; 600073; -  
DR InterPro; IPR000033; -  
DR InterPro; IPR000152; -  
DR InterPro; IPR000561; -  
DR InterPro; IPR001881; -  
DR InterPro; IPR002172; -  
DR Pfam; PF00008; EGF\_13.  
DR Pfam; PF00057; ldl\_recept\_a; 36.  
DR Pfam; PF00058; ldl\_recept\_b; 37.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 9.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS01209; LDLRA\_1; 31.  
DR PROSITE; PS50068; LDLRA\_2; 36.  
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;  
KW Receptor; EGF-like domain; Signal; Polymorphism.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 4655 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED  
FT DOMAIN 26 4423 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 4424 4446 POTENTIAL.  
FT DOMAIN 4447 4655 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 26 64 LDL-RECEPTOR CLASS A 1.  
FT DOMAIN 65 105 LDL-RECEPTOR CLASS A 2.  
FT DOMAIN 106 144 LDL-RECEPTOR CLASS A 3.  
FT DOMAIN 145 181 LDL-RECEPTOR CLASS A 4.  
FT DOMAIN 182 219 LDL-RECEPTOR CLASS A 5.  
FT DOMAIN 220 258 LDL-RECEPTOR CLASS A 6.  
FT DOMAIN 264 308 LDL-RECEPTOR CLASS A 7.  
FT DOMAIN 309 347 EGF-LIKE 1.  
FT DOMAIN 348 386 EGF-LIKE 2.  
FT DOMAIN 436 477 LDL-RECEPTOR CLASS B 1.  
FT DOMAIN 479 520 LDL-RECEPTOR CLASS B 2.  
FT DOMAIN 522 567 LDL-RECEPTOR CLASS B 3.  
FT DOMAIN 569 612 LDL-RECEPTOR CLASS B 4.  
FT DOMAIN 613 653 LDL-RECEPTOR CLASS B 5.  
FT DOMAIN 659 705 EGF-LIKE 3.  
FT DOMAIN 753 794 LDL-RECEPTOR CLASS B 6.  
FT DOMAIN 796 836 LDL-RECEPTOR CLASS B 7.  
FT DOMAIN 838 880 LDL-RECEPTOR CLASS B 8.  
FT DOMAIN 882 924 LDL-RECEPTOR CLASS B 9.  
FT DOMAIN 970 1014 EGF-LIKE 4.  
FT DOMAIN 1024 1062 LDL-RECEPTOR CLASS A 8.  
FT DOMAIN 1065 1103 LDL-RECEPTOR CLASS A 9.  
FT DOMAIN 1107 1145 LDL-RECEPTOR CLASS A 10.  
FT DOMAIN 1147 1185 LDL-RECEPTOR CLASS A 11.  
FT DOMAIN 1186 1224 LDL-RECEPTOR CLASS A 12.  
FT DOMAIN 1228 1268 LDL-RECEPTOR CLASS A 13.  
FT DOMAIN 1269 1307 LDL-RECEPTOR CLASS A 14.  
FT DOMAIN 1310 1350 LDL-RECEPTOR CLASS A 15.  
FT DOMAIN 1349 1389 EGF-LIKE 5.  
FT DOMAIN 1390 1429 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 1478 1519 LDL-RECEPTOR CLASS B 10.  
FT DOMAIN 1521 1562 LDL-RECEPTOR CLASS B 11.  
FT DOMAIN 1566 1608 LDL-RECEPTOR CLASS B 12.  
FT DOMAIN 1610 1653 LDL-RECEPTOR CLASS B 13.  
FT DOMAIN 1655 1695 LDL-RECEPTOR CLASS B 14.  
FT DOMAIN 1700 1741 EGF-LIKE 7.  
FT DOMAIN 1790 1831 LDL-RECEPTOR CLASS B 15.  
FT DOMAIN 1833 1881 LDL-RECEPTOR CLASS B 16.  
FT DOMAIN 1883 1929 LDL-RECEPTOR CLASS B 17.  
FT DOMAIN 1931 1971 LDL-RECEPTOR CLASS B 18.  
FT DOMAIN 1972 2012 LDL-RECEPTOR CLASS B 19.  
FT DOMAIN 2018 2059 EGF-LIKE 8.  
FT DOMAIN 2107 2155 LDL-RECEPTOR CLASS B 20.  
FT DOMAIN 2157 2200 LDL-RECEPTOR CLASS B 21.

FT	DOMAIN	2202	2244	LDL-RECEPTOR CLASS B 22.	
FT	DOMAIN	2246	2288	LDL-RECEPTOR CLASS B 23.	
FT	DOMAIN	2290	2331	LDL-RECEPTOR CLASS B 24.	
FT	DOMAIN	2342	2383	EGF-LIKE 9.	
FT	DOMAIN	2431	2476	LDL-RECEPTOR CLASS B 25.	
FT	DOMAIN	2478	2517	LDL-RECEPTOR CLASS B 26.	
FT	DOMAIN	2519	2561	LDL-RECEPTOR CLASS B 27.	
FT	DOMAIN	2563	2603	LDL-RECEPTOR CLASS B 28.	
FT	DOMAIN	2604	2646	LDL-RECEPTOR CLASS B 29.	
FT	DOMAIN	2651	2693	EGF-LIKE 10.	
FT	DOMAIN	2698	2738	LDL-RECEPTOR CLASS A 16.	
FT	DOMAIN	2739	2777	LDL-RECEPTOR CLASS A 17.	
FT	DOMAIN	2778	2819	LDL-RECEPTOR CLASS A 18.	
FT	DOMAIN	2820	2861	LDL-RECEPTOR CLASS A 19.	
FT	DOMAIN	2862	2901	LDL-RECEPTOR CLASS A 20.	
FT	DOMAIN	2904	2945	LDL-RECEPTOR CLASS A 21.	
FT	DOMAIN	2946	2990	LDL-RECEPTOR CLASS A 22.	
FT	DOMAIN	2991	3029	LDL-RECEPTOR CLASS A 23.	
FT	DOMAIN	3030	3070	LDL-RECEPTOR CLASS A 24.	
FT	DOMAIN	3073	3110	LDL-RECEPTOR CLASS A 25.	
FT	DOMAIN	3111	3151	EGF-LIKE 11.	
FT	DOMAIN	3152	3192	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).	
FT	DOMAIN	3239	3280	LDL-RECEPTOR CLASS B 30.	
FT	DOMAIN	3282	3331	LDL-RECEPTOR CLASS B 31.	
FT	DOMAIN	3333	3375	LDL-RECEPTOR CLASS B 32.	
FT	DOMAIN	3377	3418	LDL-RECEPTOR CLASS B 33.	
FT	DOMAIN	3419	3459	LDL-RECEPTOR CLASS B 34.	
FT	DOMAIN	3465	3509	EGF-LIKE 13.	
FT	DOMAIN	3510	3550	LDL-RECEPTOR CLASS A 26.	
FT	DOMAIN	3551	3591	LDL-RECEPTOR CLASS A 27.	
FT	DOMAIN	3592	3632	LDL-RECEPTOR CLASS A 28.	
FT	DOMAIN	3633	3673	LDL-RECEPTOR CLASS A 29.	
FT	DOMAIN	3676	3716	LDL-RECEPTOR CLASS A 30.	
FT	DOMAIN	3717	3756	LDL-RECEPTOR CLASS A 31.	
FT	DOMAIN	3757	3795	LDL-RECEPTOR CLASS A 32.	
FT	DOMAIN	3796	3834	LDL-RECEPTOR CLASS A 33.	
FT	DOMAIN	3840	3880	LDL-RECEPTOR CLASS A 34.	
FT	DOMAIN	3881	3922	LDL-RECEPTOR CLASS A 35.	
FT	DOMAIN	3926	3964	LDL-RECEPTOR CLASS A 36.	
FT	DOMAIN	3966	4006	EGF-LIKE 14.	
FT	DOMAIN	4007	4048	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).	
FT	DOMAIN	4154	4195	LDL-RECEPTOR CLASS B 35.	
FT	DOMAIN	4197	4239	LDL-RECEPTOR CLASS B 36.	
FT	DOMAIN	4242	4282	LDL-RECEPTOR CLASS B 37.	
FT	DOMAIN	4330	4368	EGF-LIKE 16.	
FT	DOMAIN	4377	4411	EGF-LIKE 17.	
Query Match					2.5%; Score 123; DB 1; Length 4655;
Best Local Similarity					16.8%; Pred. No. 3.2;
Matches 189; Conservative 118; Mismatches 303; Indels 516; Gaps 54;					
QY	25	CLGFTAHRIKRAEGW	-----EEGPTVLSDSPWNTNISGCKGRCFELQE-----	68	
Db	1200	CIGVT-----NRCDGVFDCSDNSDEAGCP-----	TRPPGMCHSDEFQCCQEDGICIPNF	1247	
QY	69	---AGPPDC-----	RDNLCKSYTSCCHDFDELCLKTARGWECTK	105	
Db	1248	WECGHPDCLYGSDEHNACVPKTCPSYFHCDN-----	GNCIH-----RAWLCDR	1292	
QY	106	DR-CGEVRNEENA---	CHCSE---DCLARGDCCTNYQVCKG-----	ESHVVD-D	147
Db	1293	DNDCGMSDEKDCPTQPFRCPSWQWQCLGH-NICVNLVSVCDGIFDCPNGTDESPLCNGN	1351		
QY	148	DCEEIKA-----	AECPAGFVRPLLIIFSVDGFRASYMKKGSKVMPNIEK--	191	
Db	1352	SCSDFNGGCTHECVQEPFGAKCLCPLGLF-----	LANDSKTCEDIDECD	1395	
QY	192	-LRSCGTHSPYMRPV-----	PTKTF	211	
Db	1396	ILGSCSQHCYNMRGSRFCSCDTGYMLESdGRCKVTASESLLLVASQNKIIADSVTSQV	1455		
QY	212	PNLYTLA-TGLY---PESHGIVGNSMYDP-----	VDFATFHLR---	245	

Db	1456	HNIYSLVENGSYIVAVDFDSISGRIFWSDATQGTWSAFQNGTDRRVVDFDSSILTETIA	1515		
QY	246	---GREKF-----	NHR-----	W--WGG	257
Db	1516	IDWVGRNLYWTDYALETIEVSKIDGSHRTVLISKNLTNPRGLADPRMNEHLFWSDWGH	1575		
QY	258	QPLWITATKQG-----	VKAGTFFW--SVVIPHERRILTLRWLTLPDHERPSVYAFYSEQ	310	
Db	1576	HPRIERASMDGSMRTVIVQDKIFWPCGLTIDYPNRLLYFM-----	DSYLDYDMDF	1624	
QY	311	PDFSGHKYGFPGPEESSYGSPTPAKRPKRKVAPKRRQ-----	ERPVA	353	
Db	1625	CDYNGH-----	HRRQVIASDLIIRHPYALTLPEDSVY	1656	
QY	354	PPKRRRKTHRMDHYAAETRODKMTN---	PLREIDKIVGQLMDGLKQLKLRRCVNVIFVG	410	
Db	1657	WTDRASTRYMRANKWHGNGQSVVMYNIQWPLGIYAVHPSKQNSVNPFCASRCSHLCILS	1716		
QY	411	DHGMEDVTC-----	DRTEFL-----	425	
Db	1717	SGPHFYSCVCPGWSLSPLDLLNCLRDDQPLITVRQHIIFGISLNPEVKSNDAMVPIAG	1776		
QY	426	-----SNYLTNVDDITLVPGLTGRIRSKFSNNAKYDPKAIIANLTCKKPDQHF	473		
Db	1777	IQGLDVEFDDAEQYIYWVEN---PGEIHRVKTDGTN-----	RTVFASISVMVGPSMNL	1826	
QY	474	K-PYLKQHLPKRLHYANNRRRIEDIHLV-----	ERRWHVARK-----	PLDVYKKPS	518
Db	1827	ALDWISRNL-----	YSTNPRTQSIIEVTLHGDIRYKTKLIANDGTALGVGFPITVDPA	1881	
QY	519	GKCFQGDHGFNDKV-----	NSMQTVFVGYGPTFKYTKVPPFENIELYNVMCD	567	
Db	1882	RGKLYWSDQGTDSGVPAKIASANMDGTSVKTLFTG-----	NLE--HLECV	1924	
QY	568	LLGLKPAPNNGTHGSLNHLRLTNTFRPTMPPEEVRPNYPG----	IMYLSDFDLGCTCDD	623	
Db	1925	TLDIEE-----	QKLYWAVTGRV-----IERGNVDGTDRLMLVHQLSHPWGIAVHD	1970	
QY	624	KVEPKNKLDLNLKRLHTKSGSTEERHLLYGRPAVLYRT--	RYDILYHTDFESGYSEIFL--	679	
Db	1971	-----	SFLYYTDEQYEVIERVDRATGANKIVLRD	1999	
QY	680	----MLLWTSYTVSKQAEVSS-----	VPDHLTSCVRPDVRSVPSFSQNC	720	
Db	2000	NVPNLRLGLQVYHRRNAAESSNGCSNNMACQICLPVPGGLFSCA-----	2044		
QY	721	AYKNDKQMSYGF-LFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRLVKKYASE	779		
Db	2045	-----CATGFKLNPDNRSCSP---YNSFIVSMLSAIRGFSLELSDHSETMVPVAGQG	2094		
QY	780	RNG-----	VNVISGPIFDYDGLDHTEDKIKQY-VEGSSI-VPVPTH	819	
Db	2095	RNALHVDVDVSSGFIYWCDFSSSVASDNARRIKPKDGSLLMNIVTH	2140		

RESULT 8  
LT23\_CAEL  
ID LT23\_CAEL STANDARD; PRT; 1323 AA.  
AC P24348;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE LET-23 RECEPTOR PROTEIN-TYROSINE KINASE PRECURSOR (EC 2.7.1.112).  
GN LET-23 OR KIN-7 OR ZK1067.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91080919; PubMed=1979659;  
RA Aroian R.V., Koga M., Mendel J.E., Ohshima Y., Sternberg P.W.;  
RT \*The let-23 gene necessary for Caenorhabditis elegans vulval



RT induction encodes a tyrosine kinase of the EGF receptor subfamily.";  
RL Nature 348:693-699(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=96177760; PubMed=8604137;  
RA Sakai T., Koga M., Ohshima Y.;  
RT "Genomic structure and 5' regulatory regions of the let-23 gene in  
the nematode C. elegans.";  
RL J. Mol. Biol. 256:548-555(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Thomas K.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP MUTANTS.  
RX MEDLINE=94147981; PubMed=8313880;  
RA Aroian R.V., Les G.M., Sternberg P.W.;  
RT "Mutations in the Caenorhabditis elegans let-23 EGFR-like gene define  
elements important for cell-type specificity and function.";  
RL EMBO J. 13:360-366(1994).  
CC -!- FUNCTION: TYROSINE KINASE RECEPTOR REQUIRED FOR THE INDUCTION OF  
C.ELEGANS VULVA. POSSIBLE RECEPTOR FOR THE INDUCTIVE SIGNAL  
REQUIRED FOR VULVAL DEVELOPMENT. ACTIVATED BY LIN-3 AND ACTS BY  
WAY OF LET-60 RAS. THE LIN-3/LET-23 PAIR IS A SIMPLIFIED VERSION  
OF THE MAMMALIAN NEUREGULIN-ERBB NETWORK.  
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
PROTEIN TYROSINE PHOSPHATE.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X57767; CAA40919.1; ALT\_SEQ.  
CC EMBL; D63426; BAA09729.1; ALT\_INIT.  
CC EMBL; Z70038; CAA93882.1; -.  
CC PIR; S13422; S13422.  
CC HSSP; P11362; 1FGI.  
CC WormPep; ZK1067.1; CE03840.  
CC InterPro; IPR000494; -.  
CC InterPro; IPR000719; -.  
CC InterPro; IPR001245; -.  
CC InterPro; IPR002174; -.  
CC Pfam; PF00757; Furin-like; 1.  
CC Pfam; PF01030; Recep\_L\_domain; 2.  
CC Pfam; PF00069; pkinase; 1.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
KW Transmembrane; Glycoprotein; Receptor; Signal.  
FT SIGNAL 1 22  
FT CHAIN 23 1323  
FT DOMAIN 23 818  
FT TRANSMEM 819 841  
FT DOMAIN 842 1323  
FT DOMAIN 885 1152  
FT NP\_BIND 891 899  
FT BINDING 919 919  
FT ACT\_SITE 1010 1010  
FT CARBOHYD 91 91  
FT CARBOHYD 169 169  
FT CARBOHYD 255 255  
FT CARBOHYD 376 376  
FT CARBOHYD 561 561  
FT CARBOHYD 655 655

FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 368 368 C -> Y (IN SY10).  
FT VARIANT 469 469 G -> R (IN MN216).  
FT VARIANT 700 700 C -> W (IN MN23).  
FT VARIANT 753 753 C -> Y (IN SY11).  
FT VARIANT 1065 1065 T -> I (IN SY16).  
FT VARIANT 1074 1074 G -> E (IN SY7).  
SQ SEQUENCE 1323 AA; 150510 MW; 6B0307EE53EEFA99 CRC64;  
  
Query Match 2.4%; Score 122; DB 1; Length 1323;  
Best Local Similarity 18.4%; Pred. No. 0.73;  
Matches 149; Conservative 106; Mismatches 290; Indels 264; Gaps 43;  
  
QY 48 SDSPTNISGSK-----GRCFELQEAAGPPDCRDNLC-----KSYTSCC 87  
Db 225 NDKCWGSGDNDQCQVYRSVCPKSCQCFYSNSTSSYEC-CDsACLGCTGHGPKNCIACS 283  
QY 88 -HDFDELCLKTARGWECTKDRCEV-----RNEENACHCEDC----LARGDCCTNYQVV 137  
Db 284 KYELDGICITCPSRKIFNHKTKGLVFNPDGRYQNGHCVKCEPPELLIENDVCVRH--- 340  
QY 138 CKGESHW--VDD--DCEEIKAACCPAGFVRPPLIIFSVDGFRASYMKKSKVMPNIEKLR 193  
Db 341 CSDGHHYDATKDRECEKCRSSCPK-----ICTVDGHLT-----NETLKNLEGCE 386  
QY 194 SCGTHSPYMRPVYPTKTFPNLYTLAT-GLYPESHGIVGNSMVDVDFDATFH-LRGREKFN 251  
Db 387 QIDGH----LIIEHAFTYEQKLVETVKIVSEYITIVQQNFYDLKFLKNLQIIEGRKLHN 442  
QY 252 HRWGG-----QPLWITATKQGVKAGTFFWSVVIPIHERILLTLRLWTLPLDHERPSVY 304  
Db 443 VRWALAIYQDDELEELSLNSKL-IKTG---AVLIMKNHRLC-----480  
QY 305 AFYSEQPDFSGHKYGPFGPEESSYSGSPFTPAKRPKRKVAPKRQRPVAPPKKRRRKIHR 364  
Db 481 --YVSKIDWSS-----IITSKGNKPSLAIAENR-----508  
QY 365 MDHYAAETRODKMTNPLREIDKIVQQLMDGLKQLKLRRCNVNIFVGDHGMEDVTCDRTEF 424  
Db 509 -DSKLCETEQ-----RVCDK-----NCNKRGCWGKEPEDCLECKT 542  
QY 425 LSNYLTNVDDITLVPGTGLGRIRSKFSNNA-KYDPKAIAN-----LTCKKPDQHFQPY 476  
Db 543 WKSVCCTVEKC---DTKGFLRNQTSMKCERCSPECETCNGELGELDCLTCRH----- 590  
QY 477 LKQHLPKRLHYAN-NRRIEDIHLLVERRWHVARKPL--DVYKPKSGKCFQGDHGFNDKV 533  
Db 591 -----KTLYNSDFGNRMCEVH-----DCPVSHFPTQKNVCEKCHPTCYDNGCTGPDSN- 638  
QY 534 NSMQTFVVGYG--PTFKYKTKVPPFENIELY-----NVMC---DL-----568  
Db 639 -----LGYGCKCKQKYAVK---YENDTIFCLOSSGMNNVCVENDLPNYYIISTYDTGV 688  
QY 569 -----LGLKPAPNNGTHGSLNHLRLTNT-FRPTMPEEVTTPNYPGIMYLSQDFD-- 616  
Db 689 IETHCEKCSIACKTCSAGRNVVQNKVCVKHVEYQPNPSEICMDQCPVNSFMVPTNNT 748  
QY 617 -----LGC-TCDD-----KVEPKNKLDLKNRLHTKSGSTERHLLY 651  
Db 749 VCKKCHHECDQNYHCANGSTGCQCKKNFTVFKGDIACQVSECCKNLPFSNPANGECLDY 808  
QY 652 GRPAVLYRTRYDILYHTDFESGYSEIFLMLL---WTSYTVSKQAESVSSVPDHLTSCVRPD 708  
Db 809 DIASRQRKTRMVIIGSVLF--GFVWMLFILLVYWRQCRIGRIGKLLKIAEMVD-MPELTPID 865  
QY 709 VRVSPFSFQNC-----LAYKNDKQMSYG 731  
Db 866 ASVRPNMSRICLIPSSSELOTKLDKLLGAG 894  
  
RESULT 9

YDT2\_SCHPO  
ID YDT2\_SCHPO STANDARD; PRT; 1888 AA.  
AC O14207;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE HYPOTHETICAL 217.4 KDA PROTEIN C6B12.02C IN CHROMOSOME I.  
GN SPAC6B12.02C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z98531; CAB11064.1; -.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 697 717 POTENTIAL.  
FT TRANSMEM 755 775 POTENTIAL.  
FT TRANSMEM 866 886 POTENTIAL.  
FT TRANSMEM 915 935 POTENTIAL.  
FT TRANSMEM 977 997 POTENTIAL.  
FT TRANSMEM 1033 1053 POTENTIAL.  
FT TRANSMEM 1336 1356 POTENTIAL.  
FT TRANSMEM 1645 1665 POTENTIAL.  
FT DOMAIN 1662 1665 POLY-LEU.  
SQ SEQUENCE 1888 AA; 217432 MW; 8AD3BBCE32397C29 CRC64;  
  
Query Match 2.4%; Score 121; DB 1; Length 1888;  
Best Local Similarity 17.2%; Pred. No. 1.4;  
Matches 123; Conservative 101; Mismatches 223; Indels 268; Gaps 32;  
  
QY 327 SYGSPFTPAKRPKRKVAPK-----RRQRPVAPPKRRRKIHRMDHYAAETRODKMTNPL 381  
Db 441 SYG-----KKAPKFLRVFARSSSHIPKMIRKRQMSKKYFSDKESDRQV--- 486  
  
QY 382 REIDKIVQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNYLTNVDDITLVPGT 441  
Db 487 --IDQVLSDWYSGKHELVOQ-----SHSYKKPSDSKSVGGN 520  
  
QY 442 LGRIRSKFSNNAKYDPKAIANLTCKKPDQHFKPYLKQHLPKR-----LHYANNR 491  
Db 521 I-----FSVNSKKHNVNINAKTAANGLSHLQNFSELLKKRKLFSLSFNVSYYKSK 574  
  
QY 492 RIEDIHLVRRWHVARKPLDVY-----KKP 517  
Db 575 KLKRTHTVHDKCQKVAK--LDHYIRDNIELNSKEREDYEGTLAVPQVNTAIRKSSRKQ 632  
  
QY 518 SGKCFQGDHGFNDKNVSMQTVFVGYPFTFKYKTKVPFENIELYNVCMCDLLGLKPA--- 574  
Db 633 KAQRFIRD--FD-----TVFFQSSSNPNYFTDVNPFNIGIWSTTFNITFRPGLSL 683  
  
QY 575 PNNG---THGSLNHLR-----TNTFRPTM-----PEEVTRPNYP 606  
Db 684 PNNFSIKTQG-LNSILQLDIVTHPFSKVYAFSCLFNIQDDVFKTEKLDKDTFTVLENLP 742  
  
QY 607 -----GIMYQSDF-DLGCTCDDKVEPKNKLDLNLKR 637  
.....

QY 638 LHTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLW-TSYTVSKQAEVSS 696  
Db 798 -----RSQVKEIAVLFQT-FQVVFYCEFGNQNNKINKVSWLASDLISKL--LSA 845  
QY 697 VPDHLTSCVRPDVRSFQNSQCLAYKNDKQMSYGLFPPYLSSSPKAKYDAFLVTNMVP 756  
Db 846 GQSGLLECYR-NLRIQAS-----DTTVIDTL-- 870  
QY 757 MYPAFKRVWNYFQRLV---KKYA-----SERNGVNVIS-----GPIF 791  
Db 871 ----FLESWSILNHLFHVYKKYALWEQVNSFFDLQKELSILEMEKIWYVIMTLNPVF 926  
QY 792 DYDYG-----LHDTEDKIKQYVEGSIPTVPTHYYSIITSCLDFTQPADKC 837  
Db 927 QIGLNGTTHSPGNNSFWPLIIRVSESAFKMHKDGHNKVVVERYLRTVFLRIHFLISEWRW 986  
QY 838 DGPLSVSSFI---LPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLD 889  
Db 987 EDVAQILFLIFDFFSHRKFNLSSEISEDPTDFDFVK-SLDRPPNL-HVTALD 1039

RESULT 10

NUCL\_CUNEE STANDARD; PRT; 252 AA.  
ID NUCL\_CUNEE  
AC P81203;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE NUCLEASE C1 (EC 3.1.30.-).  
GN NUC1CE.  
OS Cunninghamella echinulata var. echinulata.  
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales;  
OC Cunninghamellaceae; Cunninghamella.  
OX NCBI\_TaxID=76406;  
RN [1]  
RP SEQUENCE OF 15-252 FROM N.A., AND SEQUENCE OF 1-49.  
RC STRAIN=ATCC 36190;  
RX MEDLINE=98417435; PubMed=9746353;  
RA Ho H.-C., Liu F.-C., Chung J.-G., Chen L.-Y.;  
RT "Purification, characterization and complete amino acid sequence of  
RL nuclease C1 from Cunninghamella echinulata var. echinulata";  
CC -!- FUNCTION: THIS ENZYME HAS BOTH RNASE AND DNASE ACTIVITY.  
CC -!- COFACTOR: REQUIRES MAGNESIUM OR MANGANESE FOR ACTIVITY.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- SIMILARITY: BELONGS TO THE DNA/RNA NON-SPECIFIC ENDONUCLEASES  
CC FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF043517; AAC78769.2; -.  
DR InterPro; IPR001604; -.  
DR Pfam; PF01223; Endonuclease; 1.  
DR PROSITE; PS01070; NUCLEASE\_NON\_SPEC; 1.  
KW Hydrolase; Nuclease; Endonuclease; Magnesium.  
FT ACT SITE 87 87 BY SIMILARITY.  
SQ SEQUENCE 252 AA; 27498 MW; 2FA49B158AAABEC6 CRC64;

Query Match 2.4%; Score 119.5; DB 1; Length 252;  
Best Local Similarity 25.0%; Pred. No. 0.13;  
Matches 65; Conservative 33; Mismatches 113; Indels 49; Gaps 15;  
  
QY 646 ERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTS--YTVSKQAEVSSVP-DHLT 702  
.....

RT induction encodes a tyrosine kinase of the EGF receptor subfamily.";

RL Nature 348:693-699(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RX MEDLINE=96177760; PubMed=8604137;

RA Sakai T., Koga M., Ohshima Y.;

RT "Genomic structure and 5' regulatory regions of the let-23 gene in

RT the nematode C. elegans.";

RL J. Mol. Biol. 256:548-555(1996).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RA Thomas K.;

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

RN [4]

RP MUTANTS.

RX MEDLINE=94147981; PubMed=8313880;

RA Aroian R.V., Les G.M., Sternberg P.W.;

RT "Mutations in the Caenorhabditis elegans let-23 EGFR-like gene define

RT elements important for cell-type specificity and function.";

RL EMBO J. 13:360-366(1994).

CC -!- FUNCTION: TYROSINE KINASE RECEPTOR REQUIRED FOR THE INDUCTION OF

CC C.ELEGANS VULVA. POSSIBLE RECEPTOR FOR THE INDUCTIVE SIGNAL

CC REQUIRED FOR VULVAL DEVELOPMENT. ACTIVATED BY LIN-3 AND ACTS BY

CC WAY OF LET-60 RAS. THE LIN-3/LET-23 PAIR IS A SIMPLIFIED VERSION

CC OF THE MAMMALIAN NEUREGULIN-ERBB NETWORK.

CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +

CC PROTEIN TYROSINE PHOSPHATE.

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X57767; CAA40919.1; ALT\_SEQ.

CC EMBL; D63426; BAA09729.1; ALT\_INIT.

CC EMBL; Z70038; CAA93882.1; -.

CC PIR; S13422; S13422.

CC HSSP; P11362; 1FGI.

CC WormPep; ZK1067.1; CE03840.

CC InterPro; IPR000494; -.

CC InterPro; IPR000719; -.

FT	CARBOHYD	746	746	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	776	776	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	368	368	C -> Y (IN SY10).
FT	VARIANT	469	469	G -> R (IN MN216).
FT	VARIANT	700	700	C -> W (IN MN23).
FT	VARIANT	753	753	C -> Y (IN SY11).
FT	VARIANT	1065	1065	T -> I (IN SY16).
FT	VARIANT	1074	1074	G -> E (IN SY7).
SQ	SEQUENCE	1323	AA; 150510	MW; 6B0307EE53EEFA99 CRC64;

Query Match 2.4%; Score 122; DB 1; Length 1323;  
Best Local Similarity 18.4%; Pred. No. 0.73;  
Matches 149; Conservative 106; Mismatches 290; Indels 264; Gaps 43;

QY	48	SDSPWNTNISGSK-----GRCFELQEAGPPDCRCDNLC-----KSYTSCC	87
Db	225	NDKCGSGDNDQCQRVYRSVCPKSCSQCFYSNSTSSYEC-CDSACLGCTGHGPKNCIACS	283
QY	88	-HDELCCLKTARGWECTKRCGEV-----RNEENACHSEDC-----LARGDCCTNYQVV	137
Db	284	KYELDGICITCPSRKIFNHKTGRLVFNPDGRYQNGNHCVKCECPPELLIENDVCVRH---	340
QY	138	CKGESHW--VDD--DCEEIKAAECFAGFVRPPLIIFSVDGFRASYMKKSGKVMPIEKLRL	193
Db	341	CSDGHYDATKDVRECEKCRSSCPK-----ICTVDGHLT-----NETLKNLEGCE	386
QY	194	SCGTHSPYMRPVYPTKTFPNLYTLAT-GLYPESHGIVGNSMYDPVDFATFH-LRGREKFN	251
Db	387	QIDGH----LIIEHAFTYEQLKVLETVKIVSEYITIVQQNFYDLKFLKNLQIIEGRKLHN	442
QY	252	HRWGG-----QPLWITATKQGVKAGTFFWSVIPHERRILTLRWLTLPDHERPSVY	304
Db	443	VRWALAIYQCDDLEELSLNSLKL-IKTG----AVLIMKNHRLC-----	480
QY	305	AFYSEQPDFSGHKYGPFGPEESSYSGSPFPAKRPKRKVAPKRQERPVPAPKKRRRKIHR	364
Db	481	--YVSKIDWSS-----IITSKGDNKPSPSLAIAENR-----	508
QY	365	MDHYAAETRODKMTNPLREIDKIVQQLMDGLKQLKLRRCVNVIVFGDHGMEDVTCDRTEF	424
Db	509	-DSKLCETEQ-----RVCDK-----NCNKRGCWGKEPEDCLECKT	542
QY	425	LSNYLTNVDDITLVPGTLGRIRSKFSNNA-KYDPKAIAN-----LTCKKPDQHEKPY	476
Db	543	WKSVCVCEKC----DTKGFRLNQTSKMCERCSPCECTCNGLGELDLCLTCRH-----	590
QY	477	LKQHLPKRLHYAN-NRRIEDIHLVERRRWHVARKPL--DVYKKPSGKCFFQGDHGFDNKV	533











---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2001, 14:43:57 ; Search time 42.51 Seconds  
(without alignments)  
2847.779 Million cell updates/sec

Title: US-09-483-831-69  
Perfect score: 5019  
Sequence: 1 MARRSSFQSCQIISLFTFAV.....RSYPEILTLKTYLHTYESEI 915

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5019	100.0	915	4 Q13827	Q13827 homo sapien
2	4681	93.3	863	4 Q13822	Q13822 homo sapien
3	4666	93.0	863	4 Q15117	Q15117 homo sapien
4	4439.5	88.5	862	11 Q9R1E6	Q9rie6 mus musculu
5	4306	85.8	885	11 Q64610	Q64610 rattus norv
6	2084	41.5	875	11 Q63490	Q63490 rattus norv
7	2084	41.5	875	11 P97676	P97676 rattus norv
8	2081	41.5	875	11 P97675	P97675 rattus norv
9	2057.5	41.0	876	11 P70641	P70641 rattus norv
10	2040.5	40.7	875	4 O14638	O14638 homo sapien
11	1955.5	39.0	845	4 Q9NP23	Q9np23 homo sapien
12	1955.5	39.0	925	4 Q9P1P6	Q9p1p6 homo sapien
13	1151	22.9	817	14 Q90761	Q90761 fowlpox vir
14	1151	22.9	817	14 Q9J5H1	Q9j5h1 fowlpox vir
15	646.5	12.9	457	10 Q9SU82	Q9su82 arabidopsis
16	613	12.2	479	10 Q9FS13	Q9fs13 spinacia ol
17	612	12.2	496	10 Q9SU83	Q9su83 arabidopsis
18	611.5	12.2	479	10 Q42974	Q42974 oryza sativ
19	610	12.2	829	5 P90754	P90754 caenorhabdi

20	609	12.1	461	10 Q9SU81	Q9su81 arabidopsis
21	539.5	10.7	485	3 Q94323	Q94323 schizosacch
22	534	10.6	251	4 Q9NQM9	Q9nqm9 homo sapien
23	526.5	10.5	257	11 O88827	O88827 rattus norv
24	522	10.4	429	2 O69013	O69013 zymomonas m
25	512	10.2	453	4 Q9Y6X5	Q9y6x5 homo sapien
26	474.5	9.5	477	11 Q9EQG7	Q9eqg7 mus musculu
27	466.5	9.3	477	4 Q9UJA9	Q9uja9 homo sapien
28	461	9.2	433	2 Q9PAB9	Q9pab9 xylella fas
29	456.5	9.1	674	5 P90755	P90755 caenorhabdi
30	427	8.5	152	4 Q9H515	Q9h515 homo sapien
31	394.5	7.9	281	6 Q9TSB2	Q9tsb2 bos taurus
32	309.5	6.2	151	11 Q9QYV2	Q9qyv2 rattus norv
33	297.5	5.9	614	5 Q17540	Q17540 caenorhabdi
34	241.5	4.8	133	10 Q9SU80	Q9su80 arabidopsis
35	238	4.7	453	5 Q22129	Q22129 caenorhabdi
36	234	4.7	45	4 Q14555	Q14555 homo sapien
37	171	3.4	385	13 Q9PTU6	Q9ptu6 paralichthy
38	170.5	3.4	44	11 Q9RID5	Q9rid5 mus musculu
39	163	3.2	360	5 O45359	O45359 caenorhabdi
40	158.5	3.2	1054	11 Q9JMG9	Q9jmg9 mus musculu
41	145.5	2.9	465	2 O69954	O69954 streptomyce
42	142	2.8	1186	14 O55767	O55767 chilo iride
43	135	2.7	32	6 Q9TRD2	Q9trd2 bos taurus
44	134	2.7	1404	4 Q92954	Q92954 homo sapien
45	132	2.6	462	2 Q9Z4Z4	Q9z4z4 streptomyce

ALIGNMENTS

RESULT 1  
Q13827  
ID Q13827 PRELIMINARY; PRT; 915 AA.  
AC Q13827;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE AUTOTAXIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=TUMOR;  
RX MEDLINE=95074054; PubMed=7982964;  
RA Murata J., Lee H.Y., Clair T., Krutzsch H.C., Arestad A.A.,  
RA Sobel M.E., Liotta L.A., Stracke M.L.;  
RT "CDNA cloning of the human tumor motility-stimulating protein,  
autotaxin, reveals a homology with phosphodiesterases.";  
RL J. Biol. Chem. 269:30479-30484(1994).  
DR EMBL; L35594; AAA64785.1; -  
DR InterPro; IPR001212; -  
DR InterPro; IPR001604; -  
DR InterPro; IPR002591; -  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR Pfam; PF01663; Phosphodiect; 1.  
DR PRINTS; PR00022; SOMATOMEDINB.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
DR SMART; SM00477; NUC; 1.  
SQ SEQUENCE 915 AA; 105246 MW; 2BE6135CCD837780 CRC64;

Query Match 100.0%; Score 5019; DB 4; Length 915;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARRSSFQSCQIISLFTFAVGVSICLGFTAHRIKRAEGWEGPPTVLSDSPWTNIGSGCK 60  
|||||

Db 1 MARRSSFQSCQIISLFTFAVGVSICLGFTAHRIKRAEGWEGPPTVLSDSPWTNIGSGCK 60  
|||||

QY 61 GRCFELQEAQPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCEVRNEENACHC 120



```
Db 61  GRCFELQEAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCEVRNEENACHC 120
QY 121  SEDCLARGDCCTNYQVCKGESHWWDDCEEIKAACPAGEFVRPPLIIFSVDGFRASYMK 180
Db 121  SEDCLARGDCCTNYQVCKGESHWWDDCEEIKAACPAGEFVRPPLIIFSVDGFRASYMK 180
QY 181  KGSKVPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240
Db 181  KGSKVPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240
QY 241  TFHLRGREKFNHRWWGGQPLWITATKQGVKAGTFFWSVVIPHERRILTLRWLTLPDHER 300
Db 241  TFHLRGREKFNHRWWGGQPLWITATKQGVKAGTFFWSVVIPHERRILTLRWLTLPDHER 300
QY 301  PSVYAFYSEQPDFSGHKYGPFGPESSYGSPTPAKRPKRKVAPKRRQERPVAPPKKRRR 360
Db 301  PSVYAFYSEQPDFSGHKYGPFGPESSYGSPTPAKRPKRKVAPKRRQERPVAPPKKRRR 360
QY 361  KIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCD 420
Db 361  KIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCD 420
QY 421  RTEFLSNYLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKPDQHFKPYLKQH 480
Db 421  RTEFLSNYLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKPDQHFKPYLKQH 480
QY 481  LPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGCKFFQGDHGFNDKNVSMQTVE 540
Db 481  LPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGCKFFQGDHGFNDKNVSMQTVE 540
QY 541  VGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEV 600
Db 541  VGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEV 600
QY 601  TRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEEHLLYGRPAVLYRT 660
Db 601  TRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEEHLLYGRPAVLYRT 660
QY 661  RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRSFSQNCL 720
Db 661  RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRSFSQNCL 720
QY 721  AYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRLVKKYASER 780
Db 721  AYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRLVKKYASER 780
QY 781  NGVNVISGPIFDYDGLDHTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGP 840
Db 781  NGVNVISGPIFDYDGLDHTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGP 840
QY 841  LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSSYPE 900
Db 841  LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSSYPE 900
QY 901  ILTLKTYLHTYESEI 915
Db 901  ILTLKTYLHTYESEI 915
```

```
RESULT 2
Q13822
ID Q13822 PRELIMINARY; PRT; 863 AA.
AC Q13822;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE AUTOTAXIN-T.
GN ATX-T.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=95074054; PubMed=7982964;
RA Murata J., Lee H.Y., Clair T., Krutzsch H.C., Arestad A.A.,
RA Sobel M.E., Liotta L.A., Stracke M.L.;
RT "cDNA cloning of the human tumor motility-stimulating protein,
RL autotaxin, reveals a homology with phosphodiesterases.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=96158950; PubMed=8579579;
RA Lee H.Y., Murata J., Clair T., Polymeropoulos M.H., Torres R.,
RA Manrow R.E., Liotta L.A., Stracke M.L.;
RT "Cloning, chromosomal localization, and tissue expression of autotaxin
RL from human teratocarcinoma cells.";
DR EMBL; L46720; AAB00855.1; -.
DR InterPro; IPR001212; -.
DR InterPro; IPR001604; -.
DR InterPro; IPR002591; -.
DR Pfam; PF01033; Somatomedin_B; 2.
DR Pfam; PF01663; Phosphodiect; 1.
DR PRINTS; PR00022; SOMATOMEDINB.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
DR SMART; SM00477; NUC; 1.
SQ SEQUENCE 863 AA; 99017 MW; F960DCE51663BABC CRC64;
```

```
Query Match 93.3%; Score 4681; DB 4; Length 863;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 858; Conservative 2; Mismatches 3; Indels 52; Gaps 1;

QY 1 MARRSSFQSQIISLFTFAVGVISICLGTAAHRIKRAEGWEGPPTVLSDSPWTNISGSK 60
Db 1 MARRSSFQSQIISLFTFAVGVNICLGTAAHRIKRAEGWEGPPTVLSDSPWTNISGSK 60
QY 61 GRCFELQEAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCEVRNEENACHC 120
Db 61 GRCFELQEAGPPDCRCDNLCKSYTSCCHDFDELCLKTARAWECTKDRCEVRNEENACHC 120
QY 121 SEDCLARGDCCTNYQVCKGESHWWDDCEEIKAACPAGEFVRPPLIIFSVDGFRASYMK 180
Db 121 SEDCLARGDCCTNYQVCKGESHWWDDCEEIKAACPAGEFVRPPLIIFSVDGFRASYMK 180
QY 181 KGSKVPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240
Db 181 KGSKVPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240
QY 241 TFHLRGREKFNHRWWGGQPLWITATKQGVKAGTFFWSVVIPHERRILTLRWLTLPDHER 300
Db 241 TFHLRGREKFNHRWWGGQPLWITATKQGVKAGTFFWSVVIPHERRILTLRWLTLPDHER 300
QY 301 PSVYAFYSEQPDFSGHKYGPFGPESSYGSPTPAKRPKRKVAPKRRQERPVAPPKKRRR 360
Db 301 PSVYAFYSEQPDFSGHKYGPFGPE-----MTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCD 324
QY 361 KIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCD 420
Db 325 -----MTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCD 368
QY 421 RTEFLSNYLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKPDQHFKPYLKQH 480
Db 369 RTEFLSNYLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKPDQHFKPYLKQH 428
QY 481 LPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGCKFFQGDHGFNDKNVSMQTVE 540
Db 429 LPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGCKFFQGDHGFNDKNVSMQTVE 488
QY 541 VGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEV 600
Db 489 VGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEV 548
```

QY 601 TRNYPGIMYLSQDFDLGCTCDDKVEPKNKLDELNKRHLTKGSTEEHLLYGRPAVLRT 660  
Db 549 TRNYPGIMYLSQDFDLGCTCDDKVEPKNKLDELNKRHLTKGSTEEHLLYGRPAVLRT 608  
QY 661 RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVDPDLTSCVRPDVRSFSQNC 720  
Db 609 RYDILYHTDFESGYSEIFLMLPLWTSYTVSKQAEVSSVDPDLTSCVRPDVRSFSQNC 668  
QY 721 AYKNDKQMSYGFLPPYLSSSPEAKYDAFLVTNMVMPYPAKRVWNYFQRLVVKKYASER 780  
Db 669 AYKNDKQMSYGFLPPYLSSSPEAKYDAFLVTNMVMPYPAKRVWNYFQRLVVKKYASER 728  
QY 781 NGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSLDFTQPADKCDGP 840  
Db 729 NGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSLDFTQPADKCDGP 788  
QY 841 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVTDIEHLTSLDFFRKTSSRYPE 900  
Db 789 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVTDIEHLTSLDFFRKTSSRYPE 848  
QY 901 ILTLKTYLHTYESEI 915  
Db 849 ILTLKTYLHTYESEI 863  
RESULT 3  
Q15117 ID Q15117 PRELIMINARY; PRT; 863 AA.  
AC Q15117;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE PHOSPHODIESTERASE I ALPHA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96163899; PubMed=8586446;  
RA Kawagoe H., Soma O., Goji J., Nishimura N., Narita M., Inazawa J.,  
RA Nakamura H., Sano K.;  
RT "Molecular cloning and chromosomal assignment of the human brain-type  
phosphodiesterase I/nucleotide pyrophosphatase gene (PDNP2).";  
RL Genomics 30:380-384(1995).  
DR EMBL; D45421; BAA08260.1; -.  
DR InterPro; IPR001212; -.  
DR InterPro; IPR001604; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR Pfam; PF01663; Phosphodiester; 1.  
DR PRINTS; PR00022; SOMATOMEDINB.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
DR SMART; SM00477; NUC; 1.  
SQ SEQUENCE 863 AA; 99040 MW; 318EA28B1CEA7A55 CRC64;

Query Match 93.0%; Score 4666; DB 4; Length 863;  
Best Local Similarity 93.4%; Pred. NO. 0;  
Matches 855; Conservative 3; Mismatches 5; Indels 52; Gaps 1;

QY 1 MARRSSFQSQIISLFTFAVGVSICLGFTAHRIKRAEGWEGPPTVLSDSPWTNIGSCK 60  
Db 1 MARRSSFQSQIISLFTFAVGVSICLGFTAHRIKRAEGWEGPPTVLSDSPWTNIGSCK 60  
QY 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVNRNEENACHC 120  
Db 61 GRCFELQEAAGPPHCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVNRNEENACHC 120  
QY 121 SEDCLARGDCCTNYQVYCKGESHVDDDCCEEIKAAECFAGFVRPPLIIFSVDFGRASYMK 180  
Db 121 SEDCLARGDCCTNYQVYCKGESHVDDDCCEEIKAAECFAGFVRPPLIIFSVDFGRASYMK 180

QY 181 KSKVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240  
Db 181 KSKVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDA 240  
QY 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPIHERRILTLRLWLTLPDHER 300  
Db 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPIHERRILTLRLWLTLPDHER 300  
QY 301 PSVYAFYSEQPDFSGHKYGPFGPEESSYGSPTPAKRPKRKVAPKRRQERPVAPPKKRRR 360  
Db 301 PSVYAFYSEQPDFSGHKYGPFGPE----- 324  
QY 361 KIHMDHYAAETRODKMTNPLREIDKIVQQLMDGLKQLKLRRCNVNIFVGDHGMEDVTCD 420  
Db 325 -----MTNPLREIDKIVQQLMDGLKQLKLRRCNVNIFVGDHGMEDVTCD 368  
QY 421 RTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFEPYLKQH 480  
Db 369 RTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFEPYLKQH 428  
QY 481 LPKRLHYANNRRIEDIHLVERRWHVARKPLDVYKKPSGCKFFQGDHGFNDKVNMSQTVF 540  
Db 429 LPKRLHYANNRRIEDIHLVERRWHVARKPLDVYKKPSGCKFFQGDHGFNDKVNMSQTVF 488  
QY 541 VGYGPTFKYKTKVPPFENIELYNVCMCDLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEV 600  
Db 489 VGYGPTFKYKTKVPPFENIELYNVCMCDLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEV 548  
QY 601 TRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDELNKRHLTKGSTEEHLLYGRPAVLRT 660  
Db 549 TRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDELNKRHLTKGSTEEHLLYGRPAVLRT 608  
QY 661 RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVDPDLTSCVRPDVRSFSQNC 720  
Db 609 RYDILYHTDFESGYSEIFLMLPLWTSYTVSKQAEVSRVDPDLTSCVRPDVRSFSQNC 668  
QY 721 AYKNDKQMSYGFLPPYLSSSPEAKYDAFLVTNMVMPYPAKRVWNYFQRLVVKKYASER 780  
Db 669 AYKNDKQMSYGFLPPYLSSSPEAKYDAFLVTNMVMPYPAKRVWNYFQRLVVKKYASER 728  
QY 781 NGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSLDFTQPADKCDGP 840  
Db 729 NGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSLDFTQPADKCDGP 788  
QY 841 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVTDIEHLTSLDFFRKTSSRYPE 900  
Db 789 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVTDIEHLTSLDFFRKTSSRYPE 848  
QY 901 ILTLKTYLHTYESEI 915  
Db 849 ILTLKTYLHTYESEI 863  
RESULT 4  
Q9R1E6 ID Q9R1E6 PRELIMINARY; PRT; 862 AA.  
AC Q9R1E6;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE PHOSPHODIESTERASE I/NUCLEOTIDE PYROPHOSPHATASE (EC 3.1.4.1).  
GN NPPS2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57B/6;  
RA Sano K., Piao J.-H.;  
RT "Cloning and chromosomal assignment of mouse phosphodiesterase  
I/nucleotide pyrophosphatase (PD-I alpha/autotaxin).";





Db 322 -----MTNPLREIDKTVGQLDGLKQLRLHRCVNVIFVGDHG 358  
QY 414 MEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHF 473  
Db 359 MEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRAKSINNISKYDPKTIIANLTCKKPDQHF 418  
QY 474 KPYLKQHLKRLHYANNRRIEDIHLVRRWHVARKPLDVYKKPSGKCFQGDHGFNDKV 533  
Db 419 KPYMKQHLKRLHYANNRRIEDIHLVDRRHHVARKPLDVYKKPSGKCFQGDHGFNDKV 478  
QY 534 NSMQTVFVGYGPTFKYTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRLTNTFR 593  
Db 479 NSMQTVFVGYGPTFKYTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRLTNTFR 538  
QY 594 PTMPEEVRPNYPGIMYLQSDFDLCTCDDKVEPKNKLDLNLKRLHTKGSTE----- 645  
Db 539 PTMPDEVSRPNYPGIMYLQSEFDLCTCDDKVEPKNKLEELNKRLHTKGSTEAGTKFRG 598  
QY 646 -----ERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLLWTSYTV 688  
Db 599 SKHENKKNLNGSVEPRKERHLLYGRPAVLYRTSYDILYHTDFESGYSEIFLPLWTSYTI 658  
QY 689 SKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCLAYKNDKQMSYGFLPPYLSSSSPEAKYDA 748  
Db 659 SKQAEVSSIPEHLTNCVRPDVVRVSPGFSQNCLAYKNDKQMSYGFLPPYLSSSSPEAKYDA 718  
QY 749 FLVTNMVPMYPAPKRVWNYFQRLVKKYASERNGVNVISGPIFDYDGLHDTEDKIKQY 808  
Db 719 FLVTNMVPMYPAPKRVWAYFQRLVKKYASERNGVNVISGPIFDYDGLRDTDEDEIKQY 778  
QY 809 VEGSSIPVPTHYISITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWV 868  
Db 779 VEGSSIPVPTHYISITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWV 838  
QY 869 EELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915  
Db 839 EELMKMHTARVRDIEHLTGLDFYRKTSRSYSEILTLKTYLHTYESEI 885  
RESULT 6  
Q63490 PRELIMINARY; PRT; 875 AA.  
ID Q63490  
AC Q63490  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE RB13-6 ANTIGEN.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAQUE-DAWLEY; TISSUE=BRAIN;  
RX MEDLINE=95247775; PubMed=7730366;  
RA Deissler H., Lottspeich F., Rajewsky M.F.;  
RT "Affinity purification and cDNA cloning of rat neural differentiation  
RT and tumor cell surface antigen gp130RB13-6 reveals relationship to  
RT human and murine PC-1";  
RL J. Biol. Chem. 270:9849-9855(1995).  
DR EMBL; Z47987; CAA88029.1; -.  
DR InterPro; IPR001212; -.  
DR InterPro; IPR001604; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR Pfam; PF01663; Phosphodiester; 1.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; UNKNOWN\_2.  
DR SMART; SM00477; NUC; 1.  
SQ SEQUENCE 875 AA; 99087 MW; 9BA9BFBEF82DFBA6 CRC64;

Query Match

41.5%; Score 2084; DB 11; Length 875;

Best Local Similarity 43.5%; Pred. No. 9.6e-162;  
Matches 386; Conservative 160; Mismatches 248; Indels 94; Gaps 11;  
QY 57 GSKKGRCFELQEAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEEN 116  
Db 53 GSKRKKCFDSSHRGLEGCRCDSGCTDRGDCWDFEDCVKSTQIWTCSNFRCGEETRLEAA 112  
QY 117 ACHCEDCLARGDCCTNYQVCKGESHVWDDDCCEEIKAAECPAGFVRPPLIIFSVDGFRA 176  
Db 113 LCSCADDCLQRKDCCTDYKAVCQGEVPWVTEACASSQEPQCEGFDQPPVILFSDMGFRA 172  
QY 177 SYMKKGSKVPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDP 236  
Db 173 EYLQTWSTLLPNINKLKTCLGLHSKMYRAMYPTKTFPNHYTIVTGLYPESHGIIIDNNMYDV 232  
QY 237 VFDATEHLRGREKFNHRWGGQPLWITATKQGVKAGTFEFW-----SV 278  
Db 233 YLNKNFSLSSVEKSNPAWWSGQPIWLTAMYQGLKAASYWLGSADVAVNGSFPNIYRYSN 292  
QY 279 VIPHERRILTLRWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPTPAKRP 338  
Db 293 SVPYESRIATLLQWLDLPKAERPSFTYIYVEEPSAGHKSGPV-----SAG----- 338  
QY 339 KRKVAPKRRQERPVPAPKKRRRKIHRMDHYAAETRQDKWNTNPLREIDKIVGQMLDGLKQL 398  
Db 339 -----VIKALQLVDDAFGMLMEGLKQR 360  
QY 399 KLRRCVNVIFVGDHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSK-----FSNN 452  
Db 361 NLHNCVNIIVLADHGMQDTSRVEYMTDYFPEI-NFYMYQGPAPRIRTRNIPQDFETFN 419  
QY 453 AKYDPKAIIANLTCKKPDQHFKPYLKQHLKRLHYANNRRIEDIHLVRRWHVARKPLD 512  
Db 420 S-----EEIVRDLSCRKSDQHFKPYLTPLDLPKRLHYAKNVRIDKVHLMVDRQLAYR---- 471  
QY 513 VYKKPSGKCFQGDHGFNDKVNMQTVFVGYGPTFKYTKVPPFENIELYNVMDLLGLK 572  
Db 472 --NKGSSNC-EGTHGYNNEFKSMEAIFLAHGSFKEKTVIEPFENIEVYNLLCDLLHIQ 528  
QY 573 PAPNNGTHGSLNHLRLTNTFRPTMPEEVRPNYPGIMYLQSDFDLCTCDDKVEPKNKLD 632  
Db 529 PAPNNGSHGSLNHLKAPFYQPSHAEELSKSACGFTPLPKDSLNCSC-LALQTSQEE 587  
QY 633 ELNKLRLHTKG-----STEERHLLYGRPAVLYRTR-YDILYHTDFESGYSEIFLMLWTSYT 687  
Db 588 QVNQRLNLSGGEVSATEKTNLPGFRPRVIOQKNKDCLLYHREYVSGFGKAMKMPMWSSYT 647  
QY 688 VSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNCLAYKNDKQMSYGFLPPYLSSSPEAKYD 747  
Db 648 VPKPGDTSLLPPTVPDCLRADVRVDPSESQKCSFYLDQNLIDHGFLYPPAIGKNESQYD 707  
QY 748 AFLVTNMVPMYPAPKRVWNYFQRLVKKYASERNGVNVISGPIFDYDGLHDTEDKIKQ 807  
Db 708 ALITSNLVPMYKEFKKMWDFHKVLLIKYAIERNGVNVVSGPIFDYNYDGHFADPEITN 767  
QY 808 YVEGSSIPVPTHYISITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKW 867  
Db 768 YVAGTDVPVPPTHYFVVLTSCKNKTHTPDCPGWLDVLPFVVPHRPTNVESCPENKAEDLW 827  
QY 868 VEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915  
Db 828 VEERFKAHIAVRDVELLTGLDFYQEKTPQVSEILQLKTYLPTFETII 875  
RESULT 7  
P97676  
ID P97676  
AC P97676;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE ALKALINE PHOSPHODIESTERASE.  
OS Rattus norvegicus (Rat).





Db 339 -----VIKALQLVDDAFGLMEGLKQR 360

QY 399 KLRRVCNVIVFGDHGMEDVTCDRTEFLSNLTNVDDITLVPGTGLGRIRSK-----FSNN 452

Db 361 NLHNCVNIIVLADHGMDOQTCDRVEYMTDYFPEI-NFYMYQGAPAPRIRTRNIPQDFFTFN 419

QY 453 AKYDPKAIIANLTCKKPDQHFKPYPYLKQHLPKRLHYANNRRIEDIHLVRRWHVARKPLD 512

Db 420 S----EEIVRDLSCRKSDQHFKPYPYLPDLPKRLHYAKNVRIDKVLHMDVDRQWLAYR---- 471

QY 513 VYKKPSGKCFQGDHGFNDKVNMQTVFVGYGPTFKYKTKVPPFENIELYNVMDLLGLK 572

Db 472 --NKGSSNC-EGGTHGYNNEFKSMEAIFLAHGSPFKEKTVIEPFENIEVYNLLCDLLHIQ 528

QY 573 PAPNNGTHGSLNHLRLTRNTPRPTMPEEVRPNYPGIMYQLQSDFDLCTCDDKVEPKNKLD 632

Db 529 PAPNNGSHGSLNHLKAPFYQPSHAHEELSKSAGCGFTTLPKDSLNCSC-LALQTSQEE 587

QY 633 ELNKRHLTK-----GSTEERHLLYGRPAVLYRTR-YDILYHTDFESGYSEIFLMLLWTSYT 687

Db 588 QVNQRLNLRGEVSATEKTNLPFGRPRVIOKNKDCHLLYHREYVSGFGKAMKMPMWSSYT 647

QY 688 VSKQAEVSSVPDHLTSCVRPDVRVSPFSQNCLAYKNDKOMSYGFLFPPYLSSSPEAKYD 747

Db 648 VPKPGDTSSLPTVPDCLRADVRVDPSESQKCSFYLADQNIDHGFLYPPAIGNNESQYD 707

QY 748 AFLVTNMVMPYPAFKRVNNYQFQRLVVKKYASERNGVNIVSGPIFDYDYDGLHDTEDKIQ 807

Db 708 ALITSNLVPMYKEFKKMDYFHKVLLIKYAIERNGVNVVSGPIFDYNYDGHFDAPDEITN 767

QY 808 YVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSEIFLPHRPDNEESCNSSEDESKW 867

Db 768 YVAGTDVPVPTHYFVVLTSCKNKHTHTPDSCPGWLDVLPFVVPHRPTNVESCPENKAEDLW 827

QY 868 VEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915

Db 828 VEERFKAHIAVRDVELLTGLDFYQEKTPQVSEILQLKTYLPTFETII 875

RESULT 9

P70641 PRELIMINARY; PRT; 876 AA.

AC P70641;

DT 01-FEB-1997 (TReMBLrel. 02, Created)

DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE PHOSPHODIESTERASE I.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SMALL INTESTINE, PROXIMAL;

RA Sano K.;

RT "Molecular cloning of phosphodiesterase I cDNA from rat small intestine.";

RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL; D30649; BAA06333.1; -.

DR InterPro; IPR001212; -.

DR InterPro; IPR001604; -.

DR InterPro; IPR002591; -.

DR Pfam; PF01033; Somatomedin\_B; 2.

DR Pfam; PF01663; Phosphodiester; 1.

DR PROSITE; PS00524; SOMATOMEDIN\_B; UNKNOWN\_2.

DR SMART; SM00477; NUC; 1.

SQ SEQUENCE 876 AA; 99328 MW; D2F772C34A0C437A CRC64;

Query Match 41.0%; Score 2057.5; DB 11; Length 876;

Best Local Similarity 43.2%; Pred. No. 1.4e-159;

Matches 384; Conservative 159; Mismatches 251; Indels 95; Gaps 12;

QY 57 GSCKGRCFELQEAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEEN 116

Db 53 GSCRKKCFDSSHRGLEGCRCDSGCTDRGCCWDFETCVKSTQIWTCSNFRCGETRLETA 112

QY 117 ACHCEDCLARGDCCTNYQVVKGESHVDDCEETKAAECAPAGFVRPPLIIFSVDGFRA 176

Db 113 LCSCADDCLQRKDCCTDYKAVCQGEVPWVTEACASSQEPQCEGFDQPPVILFSMDGFRA 172

QY 177 SYMKKGSVMPNIEKLRSCTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDP 236

Db 173 EYLQTWSTLLPNINKLTCGLHSKYMRVYPTKTFPNHYTIVTGLYPESHGIIDNNMYDV 232

QY 237 VFDAFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFW-----SV 278

Db 233 YLNKNFSLSSVEKSNPAWWSGQPIWLTAMYQGLKAASYWPGSDVAVNGSFNPNIYRYSN 292

QY 279 VIPHERRILTLRWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPTPAKRP 338

Db 293 SVPYESRIATLLQLDLPKAERPSFYTIYVEEPSAGHKSQPV-----SAG 338

QY 339 KRKVAPKRRQERPVPAPKRRRKIHRMDHYAAETRODKMTNPLREIDKIYQQLMDGLKQL 398

Db 339 -----VIKALQLVDDAFGLMEGLKQR 360

QY 399 KLRRVCNVIVFGDHGMEDVTCDRTEFLSNLTNVDDITLVPGTGLGRIRSK-----FSNN 452

Db 361 NLHNCVNIIVLADHGMDOQTCDRVEYMTDYFPEI-NFYMYQGAPAPRIRTRNIPQDFFTFN 419

QY 453 AKYDPKAIIANLTCKKPDQHFKPYPYLKQHLPKRLHYANNRRIEDIHLVRRWHVARKPLD 512

Db 420 S----EEIVRDLSCRKSDQHFKPYPYLPDLPKRLHYAKNVRIDKVLHMDVDRQWLAYR---- 471

QY 513 VYKKPSGKCFQGDHGFNDKVNMQTVFVGYGPTFKYKTKVPPFENIELYNVMDLLGLK 572

Db 472 --NKGVPNC-EGGTHGYNNEFKSMEAIFLAHGSPFKEKTVIEPFENIEVYNLLCDLLHIQ 528

QY 573 PAPNNGTHGSLNHLRLTRNTPRPTMPEEVRPNYPGIMYQLQSDFDLCTCDDKVEPKNKLD 632

Db 529 PAPNNGSHGSLNHLKAPFYQPSHAHEELSKSAGCGFTTLPKDSLNCSC-LALQTSQEE 587

QY 633 ELNKRHLTK-----GSTEERHLLYGRPAVLYRTR-YDILYHTDFESGYSEIFLMLLWTSYT 687

Db 588 QVNQRLNLRGEVSATEKTNLPFGRPRVIOKNKDCHLLYHREYVSGFGKAMKMPMWSSYT 647

QY 688 VSKQAEVSSVPDHLTSCVRPDVRVSPFSQNCLAYKNDKOMSYGFLFPPYLSSSPEAKYD 747

Db 648 VPKPGDTSSLPTVPDCLRADVRVDPSESQKCSFYLADQNIDHGFLYPPAIGNNESQYD 707

QY 748 AFLVTNMVMPYPAFKRVNNYQFQRLVVKKYASERNGVNIVSGPIFDYDYDGLHDTEDKIQ 807

Db 708 ALITSNLVPMYKEFKKMDYFHKVLLIKYAIERNGVNVVSGPIFDYNYDGHFDAPDEITN 767

QY 808 YVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSEIFLPHRPDNE-ESCNSSEDESK 866

Db 768 YVAGTDVPVPTHYFVVLTSCKNKHTHTPDSCPGWLDVLPFVVPHRPTKPVESCPENKAEDL 827

QY 867 WVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915

Db 828 WVEERFKAHIAVRDVELLTGLDFYQEKTPQVSEILQLKTYLPTFETII 876

RESULT 10

O14638 PRELIMINARY; PRT; 875 AA.

AC O14638;

DT 01-JAN-1998 (TReMBLrel. 05, Created)

DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE PHOSPHODIESTERASE I/NUCLEOTIDE PYROPHOSPHATASE BETA (EC 3.1.4.1)

DE (PHOSPHODIESTERASE I BETA).

GN PDNP3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;





```
Db 27 SCKGRCFE-RTFG--NCRDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSL 83
QY 118 CHCEDCLARGDCCTNYQVVCKGESHWVDDCEETKAAECPCAGFVRPPLIIFSVDGFRAS 177
Db 84 CACSDCKDKGCCINYSVVCQGEKSWVEEPCESINEPQCPAGFETPTLLFSLDGFRAE 143
QY 178 YMKKGSKVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPV 237
Db 144 YLHTWGGLLPVISKLCGTYTKNMRPVYPTKTFPNHYISIVTGLYPESHGIIIDNKMYDPK 203
QY 238 FDATEHLRGREKFNHRWGGQPLWITATKQGVKAGTFW-----SVV 279
Db 204 MNASFSLKSKEKFNPEWKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGS 263
QY 280 IPHERRILTILRWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPTPAKRPK 339
Db 264 VPFEERILAVLQWLQPKDERPHFYTLYLEEPDSSGHSYGPVSSE----- 308
QY 340 RKVAPKRRQERPVPAPPKRRRKRIHRMDHYAAETRODKMTNPLREIDKIVGQMLDGLKQLK 399
Db 309 -----VIKALQRVDGMVGMMLDGLKELN 331
QY 400 LRRCVNVIFVDHGMEDVTCDRTEFLSNLYTNVDDITLVPGTGLGRIR-----SKFSNNAK 454
Db 332 LHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKYVGAARLRPSDVPDKYYS--- 388
QY 455 YDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLVERRHHVARKPLDVK 514
Db 389 FNYEGIARNLSCREPNQHFKPYLKHLFPLKRLHFAKSDRIEPLTFYLDPOWQALNPSE-- 446
QY 515 KKPSSGKCFQGDHGFNDKNVSMQTVFVGYGPTFKYTKVPPFENIELYNVMDLGLKPA 574
Db 447 RKYCG----SGFHGSDNVFNSMQALFVGYGPGFKHGEADTFENIEVYNLMCDLLNTPA 502
QY 575 PNNGTHGSLNHLRTNTFRPTMPEEVRPNYPGIMYLSQDF-----DLGCTCDDKVEPK 628
Db 503 PNNGTHGSLNHLKPNVYTPKHPEV---HP---LVQCPFTRNPRDNLGCSNPSILP- 554
QY 629 NKLDLNLKRLHTKGSTEE--RH--LLYGRPAVLRY-TRYDILYHTDFESGYSEIFLMLLW 683
Db 555 --IEDFQTQFNLTVAEEKIHKETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLW 612
QY 684 TSYTVSKQAEVSSVPDHLTSCVRPDVRSVPSFSQNLAYKNDKQMSYGLFPPYLSSSPE 743
Db 613 TSYTVDRNDSFST--EDFSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPLNKNSS 670
QY 744 AKY-DAFLVTNMVMPYPAKRVWNYFORVLVKYASERNGVNVISGPIFDYDYGLDHTE 802
Db 671 GIYSEALLTNIVPMYQSFQVIWRYFHDTLRLKYAERNGVNVSGPVDFDYDGRCDSL 730
QY 803 DKIKQ---YVEGSSIPVPTHYYSIITSCLDFTQPADKDGPLSVSSFILPHRPDNEESC 859
Db 731 ENLRQKRRVRIRNQEIILPTHFFIVLTSCKDTSQEPLHCEN-LDTLAFILPHRTDNSESCV 789
QY 860 SSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRYSYPEILTLKTYLHTYSE 914
Db 790 HGKHDSSWVEELMLHRRARITDVEHITGLSFYQQRKEPVSDILKLTHTLPTFSQE 844

RESULT 12
Q9P1P6
ID Q9P1P6 PRELIMINARY; PRT; 925 AA.
AC Q9P1P6;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE PDNP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Bozzali M., Pizutti A., Trischitta E.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF110304; AAF36094.1; -
DR EMBL; AF110280; AAF36094.1; JOINED.
DR EMBL; AF110281; AAF36094.1; JOINED.
DR EMBL; AF110283; AAF36094.1; JOINED.
DR EMBL; AF110284; AAF36094.1; JOINED.
DR EMBL; AF110285; AAF36094.1; JOINED.
DR EMBL; AF110286; AAF36094.1; JOINED.
DR EMBL; AF110287; AAF36094.1; JOINED.
DR EMBL; AF110288; AAF36094.1; JOINED.
DR EMBL; AF110289; AAF36094.1; JOINED.
DR EMBL; AF110290; AAF36094.1; JOINED.
DR EMBL; AF110291; AAF36094.1; JOINED.
DR EMBL; AF110292; AAF36094.1; JOINED.
DR EMBL; AF110293; AAF36094.1; JOINED.
DR EMBL; AF110294; AAF36094.1; JOINED.
DR EMBL; AF110295; AAF36094.1; JOINED.
DR EMBL; AF110296; AAF36094.1; JOINED.
DR EMBL; AF110297; AAF36094.1; JOINED.
DR EMBL; AF110298; AAF36094.1; JOINED.
DR EMBL; AF110299; AAF36094.1; JOINED.
DR EMBL; AF110300; AAF36094.1; JOINED.
DR EMBL; AF110301; AAF36094.1; JOINED.
DR EMBL; AF110302; AAF36094.1; JOINED.
DR EMBL; AF110303; AAF36094.1; JOINED.
DR InterPro; IPR001212; -.
DR InterPro; IPR001604; -.
DR InterPro; IPR002591; -.
DR Pfam; PF01033; Somatomedin_B; 2.
DR Pfam; PF01663; Phosphodiester; 1.
DR PRINTS; PR00022; SOMATOMEDINB.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
DR SMART; SM00477; NUC; 1.
SQ SEQUENCE 925 AA; 104924 MW; 0ECAAA063801CAFEB CRC64;

Query Match 39.0%; Score 1955.5; DB 4; Length 925;
Best Local Similarity 42.2%; Pred. No. 3.4e-151;
Matches 378; Conservative 150; Mismatches 252; Indels 115; Gaps 19;

QY 58 SCKGRCFELQEAQPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRGGEVRNEENA 117
Db 107 SCKGRCFE-RTFG--NCRDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSL 163
QY 118 CHCEDCLARGDCCTNYQVVCKGESHWVDDCEETKAAECPCAGFVRPPLIIFSVDGFRAS 177
Db 164 CACSDCKDKGCCINYSVVCQGEKSWVEEPCESINEPQCPAGFETPTLLFSLDGFRAE 223
QY 178 YMKKGSKVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPV 237
Db 224 YLHTWGGLLPVISKLCGTYTKNMRPVYPTKTFPNHYISIVTGLYPESHGIIIDNKMYDPK 283
QY 238 FDATEHLRGREKFNHRWGGQPLWITATKQGVKAGTFW-----SVV 279
Db 284 MNASFSLKSKEKFNPEWKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGS 343
QY 280 IPHERRILTILRWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPTPAKRPK 339
Db 344 VPFEERILAVLQWLQPKDERPHFYTLYLEEPDSSGHSYGPVSSE----- 388
QY 340 RKVAPKRRQERPVPAPPKRRRKRIHRMDHYAAETRODKMTNPLREIDKIVGQMLDGLKQLK 399
Db 389 -----VIKALQRVDGMVGMMLDGLKELN 411
QY 400 LRRCVNVIFVDHGMEDVTCDRTEFLSNLYTNVDDITLVPGTGLGRIR-----SKFSNNAK 454
Db 412 LHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKYVGAARLRPSDVPDKYYS--- 468
QY 455 YDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLVERRHHVARKPLDVK 514
Db 469 FNYEGIARNLSCREPNQHFKPYLKHLFPLKRLHFAKSDRIEPLTFYLDPOWQALNPSE-- 526
```







Db 388 GYDNMFFSMRSIFVGYGPRFRRGIKVPSFENVQVYNAAVEILGLRPAPNNGS 439

Search completed: July 19, 2001, 14:46:48  
Job time: 171 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2001, 14:45:00 ; Search time 34.65 Seconds  
(without alignments)  
22.745 Million cell updates/sec

Title: US-09-483-831-69\_COPY\_201\_213  
Perfect score: 76  
Sequence: 1 YMRPVYPTKTFPN 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_0601:\*  
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:\*  
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:\*  
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:\*  
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:\*  
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:\*  
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:\*  
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:\*  
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:\*  
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:\*  
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:\*  
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:\*  
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:\*  
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:\*  
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:\*  
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:\*  
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:\*  
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:\*  
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:\*  
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:\*  
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:\*  
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:\*  
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	13	22	AA1992
2	76	100.0	788	17	AA198579
3	76	100.0	829	17	AA198578
4	76	100.0	858	22	AA19986
5	76	100.0	858	22	AA19995
6	76	100.0	858	22	AA19996
7	76	100.0	858	22	AA19997
8	76	100.0	858	22	AA19998
9	76	100.0	858	22	AA19999
10	76	100.0	859	22	AA19991
11	76	100.0	863	22	AA19988
					Autotaxin phosphod
					Autotaxin derived
					Autotaxin derived
					Rat autotaxin prot
					Rat autotaxin vari
					Rat autotaxin vari
					Rat autotaxin vari
					Rat autotaxin vari
					Human autotaxin pr
					Human teratocarcin

12	76	100.0	915	17	AA1986596	A2058 autotaxin pr
13	76	100.0	915	22	AA1971987	Human melanoma aut
14	76	100.0	979	17	AA1986580	Autotaxin derived
15	71	93.4	849	17	AA1986595	N-tera 2D1 autotax
16	70	92.1	885	22	AA1971989	Rat brain autotaxi
17	69	90.8	873	21	AA1983620	Human PC-1 polypep
18	69	90.8	873	21	AA1983621	Variant human PC-1
19	69	90.8	925	16	AA1979148	Human insulin rece
20	69	90.8	925	20	AA1993355	Insulin receptor t
21	69	90.8	925	21	AA1980195	Breast cancer prot
22	49	64.5	400	21	AA1929475	Arabidopsis thalia
23	49	64.5	457	21	AA1931321	Arabidopsis thalia
24	49	64.5	461	21	AA1929474	Arabidopsis thalia
25	49	64.5	465	21	AA1929473	Arabidopsis thalia
26	43	56.6	173	20	AA1970784	Renal cancer assoc
27	43	56.6	263	18	AA1940374	Human breast cance
28	43	56.6	263	18	AA1901573	Protein encoded by
29	43	56.6	263	21	AA1984037	Amino acid sequenc
30	43	56.6	562	21	AA1953329	Human colon cancer
31	43	56.6	579	18	AA1940378	Human breast cance
32	43	56.6	579	21	AA1984041	Amino acid sequenc
33	42	55.3	411	20	AA1934445	Porphyromonas ging
34	42	55.3	428	20	AA1934324	Porphyromonas ging
35	42	55.3	528	22	AA1979259	Corynebacterium gl
36	42	55.3	530	22	AA1979258	Corynebacterium gl
37	41	53.9	109	19	AA1952839	Secreted protein e
38	41	53.9	453	20	AA1975229	Human secreted pro
39	41	53.9	453	20	AA1902378	Polypeptide identi
40	41	53.9	453	22	AA1990545	Human secreted pro
41	40	52.6	199	21	AA1944375	Arabidopsis thalia
42	40	52.6	199	21	AA1953821	Arabidopsis thalia
43	40	52.6	199	22	AA1968971	Sheep mitosis arre
44	40	52.6	199	22	AA1968972	Human mitosis arre
45	40	52.6	205	18	AA1913020	Human MAD2 protein

ALIGNMENTS

RESULT 1	
AA1971992	
ID	AA1971992 standard; peptide; 13 AA.
XX	
AC	AA1971992;
XX	
DT	28-MAR-2001 (first entry)
XX	
DE	Autotaxin phosphodiesterase catalytic site.
XX	
KW	Autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;
KW	adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;
KW	therapy; metabolic disease; obesity; dyslipidemia; serum glucose;
KW	non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;
KW	type I phosphodiesterase activity; ATP pyrophosphatase activity;
KW	ATPase; adenosine-5'-triphosphatase activity;
KW	phosphodiesterase catalytic site.
XX	
OS	Unidentified.
XX	
PN	WO200068386-A1.
XX	
PD	16-NOV-2000.
XX	
PF	
XX	
XX	05-MAY-2000; 2000WO-US12402.
PR	
XX	
XX	07-MAY-1999; 99US-0306979.
PA	
XX	(ZYMO ) ZYMOGENETICS INC.
PI	
XX	Kelly JD;
DR	
XX	WPI; 2001-007397/01.

PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
XX obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
PS Disclosure; Page 3; 126pp; English.

XX The present sequence is autotaxin phosphodiesterase catalytic  
CC site. Phosphorylation of threonine residue at position 10 of this  
CC sequence is required for autotaxin to undergo autophosphorylation  
CC and exhibit phosphodiesterase and motility-stimulating activities.  
CC Autotaxin is a glycoprotein cytokine which increases  
CC insulin signalling in adipose tissue by producing substrate for  
CC adenosine receptors, resulting in inhibition of lipolysis, and  
CC decreased hepatic gluconeogenesis and serum glucose levels, and  
CC increased insulin sensitivity. It also inhibits differentiation  
CC of adipocytes. Autotaxin has type I phosphodiesterase, adenosine-5'-  
CC triphosphatase (ATPase) and ATP pyrophosphatase activities.  
CC Autotaxin and its analogues are used to stimulate glucose uptake  
CC by cells, particularly to reduce serum glucose levels for  
CC treatment of non-insulin dependent diabetes (NIDDM) in humans,  
CC or generally any condition associated with elevated serum levels  
CC of glucose, lipid or free fatty acid (e.g. obesity or dyslipidemia).  
CC Transgenic animals that overexpress autotaxin are models for  
CC human metabolic diseases.

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 76; DB 22; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.8e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | |  
Db 1 ymrpvpyptkftfn 13

RESULT 2  
AAR86579  
ID AAR86579 standard; Protein; 788 AA.  
AC AAR86579;

28-JUN-1996 (first entry)

Autotaxin derived from teratocarcinoma N-tera 2D1 cells.

Autotaxin; ATX; cytokine; autocrine motility stimulating protein; AMF;  
melanoma cell; tumour; antibody; cancer diagnosis; therapy.

Homo sapiens.

WO9532221-A2.

30-NOV-1995.

24-MAY-1995; 95WO-US06613.

28-NOV-1994; 94US-0346455.

25-MAY-1994; 94US-0249182.

(USSH ) US DEPT HEALTH & HUMAN SERVICES.

Krutzch J, Liotta L, Murata J, Schiffmann E, Stracke M;

WPI; 1996-020533/02.

N-PSDB; AAT06613.

Autotaxin motility stimulating protein, and DNA encoding it - used

in cancer diagnosis and therapy

Claim 4; Page 62-65; 112pp; English.

AAR86559-R86596 represent autotaxin (ATX) and fragments of it. ATX is

CC an autocrine motility stimulating protein which is present in cancer  
CC cells. ATX stimulates both random and directed migration of melanoma  
CC cells. The tumorous form of ATX is a secreted protein, while the  
CC transmembrane bound form is not present in tumour cells. The cDNA  
CC encoding this sequence can be used in a vector, to transform cells. The  
CC recombinant cells can then be used to produce the peptide sequences.  
CC Antibodies specific for these sequences can be produced, and can be used  
CC in cancer diagnosis and therapy. Different sites of localisation of the  
CC protein are utilised for diagnosis and prognosis of the stages of tumour  
CC progression. The sequences can be used in treatment methods to  
CC advantageously block the activity of the secreted form of AXT, while  
CC having little effect on the membrane form of AXT.

XX SQ Sequence 788 AA;

Query Match 100.0%; Score 76; DB 17; Length 788;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | |  
Db 126 ymrpvpyptkftfn 138

RESULT 3  
AAR86578

ID AAR86578 standard; Protein; 829 AA.

AC AAR86578;

28-JUN-1996 (first entry)

Autotaxin derived from melanoma cell line A2058.

Autotaxin; ATX; cytokine; autocrine motility stimulating protein; AMF;  
melanoma cell; tumour; antibody; cancer diagnosis; therapy.

Homo sapiens.

WO9532221-A2.

30-NOV-1995.

24-MAY-1995; 95WO-US06613.

28-NOV-1994; 94US-0346455.

25-MAY-1994; 94US-0249182.

(USSH ) US DEPT HEALTH & HUMAN SERVICES.

Krutzch J, Liotta L, Murata J, Schiffmann E, Stracke M;

WPI; 1996-020533/02.

N-PSDB; TO6612.

Autotaxin motility stimulating protein, and DNA encoding it - used

in cancer diagnosis and therapy

Claim 4; Page 57-60; 112pp; English.

AAR86559-R86596 represent autotaxin (ATX) and fragments of it. ATX is  
an autocrine motility stimulating protein which is present in cancer  
cells. ATX stimulates both random and directed migration of melanoma  
cells. The tumorous form of ATX is a secreted protein, while the  
transmembrane bound form is not present in tumour cells. The cDNA  
encoding this sequence can be used in a vector, to transform cells. The  
recombinant cells can then be used to produce the peptide sequences.  
Antibodies specific for these sequences can be produced, and can be used  
in cancer diagnosis and therapy. Different sites of localisation of the  
protein are utilised for diagnosis and prognosis of the stages of tumour  
progression. The sequences can be used in treatment methods to  
advantageously block the activity of the secreted form of AXT, while

CC having little effect on the membrane form of AXT.  
XX  
SQ Sequence 829 AA;  
  
Query Match 100.0%; Score 76; DB 17; Length 829;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | | | |  
Db 115 ymrpvpyptktfpn 127  
  
RESULT 4  
ID AAY71986 standard; Protein; 858 AA.  
XX  
AC AAY71986;  
XX  
DT 28-MAR-2001 (first entry)  
DE Rat autotaxin protein.  
XX  
KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity.  
XX  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT Domain 5..15  
FT Binding-site /note= "Putative transmembrane domain"  
FT 122..124  
FT /label= RGD\_binding\_domain  
FT 143..158  
FT /label= Epitope  
FT /note= "This region is specifically claimed in claim 10"  
FT 149..158  
FT /label= Epitope  
FT /note= "This region is specifically claimed in claim 10;  
FT This region is absent in rat brain autotaxin designated  
FT as PD-1alpha sequence (AAY71989)"  
FT 196..208  
FT /note= "Autotaxin phosphodiesterase catalytic site"  
FT 585..595  
FT /label= Epitope  
FT /note= "This region is specifically claimed in claim 10"  
XX  
PN WO200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
XX  
DR WPI; 2001-007397/01.  
DR N-PSDB; AAD02131.  
XX  
PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Claim 1; Page 101-104; 126pp; English.  
XX  
CC The present sequence is rat autotaxin protein. Autotaxin is a

CC glycoprotein cytokine which increases insulin  
CC signalling in adipose tissue by producing substrate for adenosine  
CC receptors, resulting in inhibition of lipolysis, decreased hepatic  
CC gluconeogenesis and serum glucose levels, and increased insulin  
CC sensitivity. It also inhibits differentiation of adipocytes.  
CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase  
CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its  
CC analogues are used to stimulate glucose uptake by cells, particularly  
CC to reduce serum glucose levels for treatment of non-insulin dependent  
CC diabetes (NIDDM) in humans, or generally any condition associated  
CC with elevated serum levels of glucose, lipid or free fatty acid  
CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress  
CC autotaxin are models for human metabolic diseases.  
XX  
SQ Sequence 858 AA;  
  
Query Match 100.0%; Score 76; DB 22; Length 858;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | | | |  
Db 196 ymrpvpyptktfpn 208  
  
RESULT 5  
ID AAY71995  
XX AAY71995 standard; Protein; 858 AA.  
AC AAY71995;  
XX  
DT 28-MAR-2001 (first entry)  
DE Rat autotaxin variant (A93V).  
XX  
KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; mutant; variant.  
XX  
OS Rattus sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 93 /note= "Wild type Ala substituted by Val"  
FT  
XX  
PN WO200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
XX  
DR WPI; 2001-007397/01.  
XX  
PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Disclosure; Page -; 126pp; English.  
XX  
CC The present sequence is variant (A93V) of rat autotaxin protein.  
CC Autotaxin is a glycoprotein cytokine which increases insulin  
CC signalling in adipose tissue by producing substrate for adenosine  
CC receptors, resulting in inhibition of lipolysis, decreased hepatic



CC gluconeogenesis and serum glucose levels, and increased insulin  
CC sensitivity. It also inhibits differentiation of adipocytes.  
CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase  
CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its  
CC analogues are used to stimulate glucose uptake by cells, particularly  
CC to reduce serum glucose levels for treatment of non-insulin dependent  
CC diabetes (NIDDM) in humans, or generally any condition associated  
CC with elevated serum levels of glucose, lipid or free fatty acid  
CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress  
CC autotaxin are models for human metabolic diseases.  
CC Note: The present sequence is not shown in the specification  
CC but is derived from rat autotaxin protein sequence found in  
CC page 101-104 of sequence listing (AAY71986).  
XX  
SQ Sequence 858 AA;

Query Match 100.0%; Score 76; DB 22; Length 858;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | |  
Db 196 ymrpvpyptktpn 208

RESULT 6  
AAY71996  
ID AAY71996 standard; Protein; 858 AA.  
XX  
AC AAY71996;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE Rat autotaxin variant (A194V).  
XX  
KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; mutant; variant.  
XX  
OS Rattus sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 194 /note= "Wild type Ala substituted by Val"  
FT  
XX  
PN WO200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
XX  
DR WPI; 2001-007397/01.  
XX  
PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Disclosure; Page -; 126pp; English.  
XX

CC The present sequence is variant (A194V) of rat autotaxin protein.  
CC Autotaxin is a glycoprotein cytokine which increases insulin  
CC signalling in adipose tissue by producing substrate for adenosine  
CC receptors, resulting in inhibition of lipolysis, decreased hepatic

CC gluconeogenesis and serum glucose levels, and increased insulin  
CC sensitivity. It also inhibits differentiation of adipocytes.  
CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase  
CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its  
CC analogues are used to stimulate glucose uptake by cells, particularly  
CC to reduce serum glucose levels for treatment of non-insulin dependent  
CC diabetes (NIDDM) in humans, or generally any condition associated  
CC with elevated serum levels of glucose, lipid or free fatty acid  
CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress  
CC autotaxin are models for human metabolic diseases.  
CC Note: The present sequence is not shown in the specification  
CC but is derived from rat autotaxin protein sequence found in  
CC page 101-104 of sequence listing (AAY71986).  
XX  
SQ Sequence 858 AA;

Query Match 100.0%; Score 76; DB 22; Length 858;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | |  
Db 196 ymrpvpyptktpn 208

RESULT 7  
AAY71997  
ID AAY71997 standard; Protein; 858 AA.  
XX  
AC AAY71997;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE Rat autotaxin variant (S236T).  
XX  
KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; mutant; variant.  
XX  
OS Rattus sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 236 /note= "Wild type Ser substituted by Thr"  
FT  
XX  
PN WO200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
XX  
DR WPI; 2001-007397/01.  
XX

PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Disclosure; Page -; 126pp; English.  
XX

CC The present sequence is variant (S236T) of rat autotaxin protein.  
CC Autotaxin is a glycoprotein cytokine which increases insulin  
CC signalling in adipose tissue by producing substrate for adenosine  
CC receptors, resulting in inhibition of lipolysis, decreased hepatic

CC gluconeogenesis and serum glucose levels, and increased insulin  
CC sensitivity. It also inhibits differentiation of adipocytes.  
CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase  
CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its  
CC analogues are used to stimulate glucose uptake by cells, particularly  
CC to reduce serum glucose levels for treatment of non-insulin dependent  
CC diabetes (NIDDM) in humans, or generally any condition associated  
CC with elevated serum levels of glucose, lipid or free fatty acid  
CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress  
CC autotaxin are models for human metabolic diseases.  
CC Note: The present sequence is not shown in the specification  
CC but is derived from rat autotaxin protein sequence found in  
CC page 101-104 of sequence listing (AAY71986).  
XX  
SQ Sequence 858 AA;

Query Match 100.0%; Score 76; DB 22; Length 858;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | | | |  
Db 196 ymrpvypktktpn 208

RESULT 8  
AAY71998  
ID AAY71998 standard; Protein; 858 AA.  
XX  
AC AAY71998;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE Rat autotaxin variant (R265K).  
XX  
KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; mutant; variant.  
XX  
OS Rattus sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 265 /note= "Wild type Arg substituted by Lys"  
FT  
XX  
PN WO200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
XX  
DR WPI; 2001-007397/01.  
XX  
PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Disclosure; Page -; 126pp; English.  
XX  
CC The present sequence is variant (R265K) of rat autotaxin protein.  
CC Autotaxin is a glycoprotein cytokine which increases insulin  
CC signalling in adipose tissue by producing substrate for adenosine  
CC receptors, resulting in inhibition of lipolysis, decreased hepatic

CC gluconeogenesis and serum glucose levels, and increased insulin  
CC sensitivity. It also inhibits differentiation of adipocytes.  
CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase  
CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its  
CC analogues are used to stimulate glucose uptake by cells, particularly  
CC to reduce serum glucose levels for treatment of non-insulin dependent  
CC diabetes (NIDDM) in humans, or generally any condition associated  
CC with elevated serum levels of glucose, lipid or free fatty acid  
CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress  
CC autotaxin are models for human metabolic diseases.  
CC Note: The present sequence is not shown in the specification  
CC but is derived from rat autotaxin protein sequence found in  
CC page 101-104 of sequence listing (AAY71986).  
XX  
SQ Sequence 858 AA;

Query Match 100.0%; Score 76; DB 22; Length 858;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | | | |  
Db 196 ymrpvypktktpn 208

RESULT 9  
AAY71999  
ID AAY71999 standard; Protein; 858 AA.  
XX  
AC AAY71999;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE Rat autotaxin variant (S289T).  
XX  
KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; mutant; variant.  
XX  
OS Rattus sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 289 /note= "Wild type Ser substituted by Thr"  
FT  
XX  
PN WO200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
XX  
DR WPI; 2001-007397/01.  
XX  
PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Disclosure; Page -; 126pp; English.  
XX  
CC The present sequence is variant (S289T) of rat autotaxin protein.  
CC Autotaxin is a glycoprotein cytokine which increases insulin  
CC signalling in adipose tissue by producing substrate for adenosine  
CC receptors, resulting in inhibition of lipolysis, decreased hepatic

CC gluconeogenesis and serum glucose levels, and increased insulin  
CC sensitivity. It also inhibits differentiation of adipocytes.  
CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase  
CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its  
CC analogues are used to stimulate glucose uptake by cells, particularly  
CC to reduce serum glucose levels for treatment of non-insulin dependent  
CC diabetes (NIDDM) in humans, or generally any condition associated  
CC with elevated serum levels of glucose, lipid or free fatty acid  
CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress  
CC autotaxin are models for human metabolic diseases.  
CC Note: The present sequence is not shown in the specification  
CC but is derived from rat autotaxin protein sequence found in  
CC page 101-104 of sequence listing (AAY71986).  
XX  
SQ Sequence 858 AA;

Query Match 100.0%; Score 76; DB 22; Length 858;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPPTKTFPN 13  
| | | | | | | | | | | | | | |  
Db 196 ymrpvypptktfpn 208

RESULT 10  
AAY71991

ID AAY71991 standard; Protein; 859 AA.  
XX  
AC AAY71991;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE Human autotaxin protein.  
XX  
KW Human; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 8..26  
FT /note= "Putative transmembrane domain"  
FT Cleavage-site 44..45  
FT Binding-site 123..125  
FT /label= RGD\_binding\_domain  
FT Active-site 197..209  
FT /note= "Putative phosphodiesterase active site"  
XX  
XX WO200068386-A1.  
PN  
XX  
XX 16-NOV-2000.  
PD  
XX  
XX 05-MAY-2000; 2000WO-US12402.  
PF  
XX  
XX 07-MAY-1999; 99US-0306979.  
PR  
XX  
XX (ZYMO ) ZYMOGENETICS INC.  
PA  
XX  
XX Kelly JD;  
PI  
XX  
XX WPI; 2001-007397/01.  
DR  
XX  
XX N-PSDB; AAD02133.  
DR  
XX  
XX New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
PT  
XX  
XX Claim 27; Page 119-121; 126pp; English.  
PS

XX  
CC The present sequence is human autotaxin protein.  
CC Autotaxin is a glycoprotein cytokine which increases  
CC insulin signalling in adipose tissue by producing substrate for  
CC adenosine receptors, resulting in inhibition of lipolysis,  
CC decreased hepatic gluconeogenesis and serum glucose levels, and  
CC increased insulin sensitivity. It also inhibits differentiation  
CC of adipocytes. Autotaxin has type I phosphodiesterase, adenosine-5'-  
CC triphosphatase (ATPase) and ATP pyrophosphatase activities.  
CC Autotaxin and its analogues are used to stimulate glucose uptake  
CC by cells, particularly to reduce serum glucose levels for  
CC treatment of non-insulin dependent diabetes (NIDDM) in humans,  
CC or generally any condition associated with elevated serum levels  
CC of glucose, lipid or free fatty acid (e.g. obesity or dyslipidemia).  
CC Transgenic animals that overexpress autotaxin are models for  
CC human metabolic diseases.  
XX  
SQ Sequence 859 AA;

Query Match 100.0%; Score 76; DB 22; Length 859;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPPTKTFPN 13  
| | | | | | | | | | | | | | |  
Db 197 ymrpvypptktfpn 209

RESULT 11  
AAY71988

ID AAY71988 standard; Protein; 863 AA.  
XX  
AC AAY71988;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE Human teratocarcinoma autotaxin.  
XX  
KW Human; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; teratocarcinoma.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Binding-site 127..129  
FT /label= RGD\_binding\_domain  
FT Active-site 201..213  
FT /note= "Autotaxin phosphodiesterase catalytic site"  
XX  
XX WO200068386-A1.  
PN  
XX  
XX 16-NOV-2000.  
PD  
XX  
XX 05-MAY-2000; 2000WO-US12402.  
PF  
XX  
XX 07-MAY-1999; 99US-0306979.  
PR  
XX  
XX (ZYMO ) ZYMOGENETICS INC.  
PA  
XX  
XX Kelly JD;  
PI  
XX  
XX WPI; 2001-007397/01.  
DR  
XX  
XX New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
PT  
XX  
XX Claim 27; Page 108-110; 126pp; English.  
PS  
XX

CC The present sequence is autotaxin isolated from human  
CC teratocarcinoma cells. Autotaxin is a glycoprotein cytokine which  
CC increases insulin signalling in adipose tissue by producing substrate  
CC for adenosine receptors, resulting in inhibition of lipolysis,  
CC decreased hepatic gluconeogenesis and serum glucose levels, and  
CC increased insulin sensitivity. It also inhibits differentiation  
CC of adipocytes. Autotaxin has type I phosphodiesterase, adenosine-5'-  
CC triphosphatase (ATPase) and ATP pyrophosphatase activities.  
CC Autotaxin and its analogues are used to stimulate glucose uptake  
CC by cells, particularly to reduce serum glucose levels for  
CC treatment of non-insulin dependent diabetes (NIDDM) in humans,  
CC or generally any condition associated with elevated serum levels  
CC of glucose, lipid or free fatty acid (e.g. obesity or dyslipidemia).  
CC Transgenic animals that overexpress autotaxin are models for  
CC human metabolic diseases.  
XX  
SQ Sequence 863 AA;

Query Match 100.0%; Score 76; DB 22; Length 863;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFFN 13  
|||||  
Db 201 ymrpvpyptktfnp 213

RESULT 12  
AAR86596  
ID AAR86596 standard; Protein; 915 AA.  
XX  
AC AAR86596;  
XX  
DT 01-JUL-1996 (first entry)  
XX  
DE A2058 autotaxin protein.  
XX  
KW Autotaxin; ATX; cytokine; autocrine motility stimulating protein; AMF;  
KW melanoma cell; tumour; antibody; cancer diagnosis; therapy.  
XX  
OS Homo sapiens.  
XX  
PN W09532221-A2.  
XX  
PD 30-NOV-1995.  
XX  
PF 24-MAY-1995; 95WO-US06613.  
XX  
PR 28-NOV-1994; 94US-0346455.  
PR 25-MAY-1994; 94US-0249182.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Krutzch J, Liotta L, Murata J, Schiffmann E, Stracke M;  
XX  
DR WPI; 1996-020533/02.  
XX  
PT Autotaxin motility stimulating protein, and DNA encoding it - used  
PT in cancer diagnosis and therapy  
XX  
PS Claim 4; Page 91-94; 112pp; English.  
XX  
CC AAR86559-R86596 represent autotaxin (ATX) and fragments of it. This  
CC sequence represents the full length protein sequence of the A2058  
CC melanoma cell line ATX protein. ATX is an autocrine motility  
CC stimulating protein which is present in cancer cells. ATX stimulates  
CC both random and directed migration of melanoma cells. The tumorous form  
CC of ATX is a secreted protein, while the transmembrane bound form is not  
CC present in tumour cells. The cDNA encoding this sequence can be used in  
CC a vector, to transform cells. The recombinant cells can then be used to  
CC produce the peptide sequences. Antibodies specific for these sequences  
CC can be produced, and can be used in cancer diagnosis and therapy.

CC Different sites of localisation of the protein are utilised for diagnosis  
CC and prognosis of the stages of tumour progression. The sequences can be  
CC used in treatment methods to advantageously block the activity of the  
CC secreted form of AXT, while having little effect on the membrane form of  
CC AXT.  
XX  
SQ Sequence 915 AA;

Query Match 100.0%; Score 76; DB 17; Length 915;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFFN 13  
|||||  
Db 201 ymrpvpyptktfnp 213

RESULT 13  
AAY71987  
ID AAY71987 standard; Protein; 915 AA.  
XX  
AC AAY71987;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE Human melanoma autotaxin.  
XX  
KW Human; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; melanoma.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Binding-site 127..129  
FT /label= RGD\_binding\_domain  
FT Active-site 201..213  
FT /note= "Autotaxin phosphodiesterase catalytic site"  
XX  
PN W0200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
XX  
DR WPI; 2001-007397/01.  
XX  
PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Claim 27; Page 105-108; 126pp; English.  
XX  
CC The present sequence is autotaxin isolated from human melanoma  
CC cells. Autotaxin is a glycoprotein cytokine which increases  
CC insulin signalling in adipose tissue by producing substrate for  
CC adenosine receptors, resulting in inhibition of lipolysis,  
CC decreased hepatic gluconeogenesis and serum glucose levels, and  
CC increased insulin sensitivity. It also inhibits differentiation  
CC of adipocytes. Autotaxin has type I phosphodiesterase, adenosine-5'-  
CC triphosphatase (ATPase) and ATP pyrophosphatase activities.  
CC Autotaxin and its analogues are used to stimulate glucose uptake  
CC by cells, particularly to reduce serum glucose levels for  
CC treatment of non-insulin dependent diabetes (NIDDM) in humans,





CC sequences can be produced, and can be used in cancer diagnosis and  
CC therapy. Different sites of localisation of the protein are utilised for  
CC diagnosis and prognosis of the stages of tumour progression. The  
CC sequences can be used in treatment methods to advantageously block the  
CC activity of the secreted form of AXT, while having little effect on the  
CC membrane form of AXT.

XX

SQ Sequence 849 AA;

Query Match 93.4%; Score 71; DB 17; Length 849;  
Best Local Similarity 92.3%; Pred. No. 0.00091;  
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
:|||||  
Db 201 hmrpvptktfpn 213

Search completed: July 19, 2001, 14:45:00  
Job time: 208 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 23, 2001, 13:16:07 ; Search time 20.3 Seconds  
(without alignments)  
12.901 Million cell updates/sec

Title: US-09-483-831-69\_COPY\_201\_213

Perfect score: 76  
Sequence: 1 YMRPVYPTKTFPN 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB ID	Description
	Score	Match	Length		
1	76	100.0	788	1 US-08-346-455B-36	Sequence 36, Appl
2	76	100.0	788	3 US-08-977-221-36	Sequence 36, Appl
3	76	100.0	788	5 PCT-US95-06613-36	Sequence 36, Appl
4	76	100.0	829	1 US-08-346-455B-34	Sequence 34, Appl
5	76	100.0	829	3 US-08-977-221-34	Sequence 34, Appl
6	76	100.0	829	5 PCT-US95-06613-34	Sequence 34, Appl
7	76	100.0	915	1 US-08-346-455B-69	Sequence 69, Appl
8	76	100.0	915	3 US-08-977-221-69	Sequence 69, Appl
9	76	100.0	915	5 PCT-US95-06613-69	Sequence 69, Appl
10	76	100.0	979	1 US-08-346-455B-38	Sequence 38, Appl
11	76	100.0	979	3 US-08-977-221-38	Sequence 38, Appl
12	76	100.0	979	5 PCT-US95-06613-38	Sequence 38, Appl
13	71	93.4	861	1 US-08-346-455B-67	Sequence 67, Appl
14	71	93.4	861	3 US-08-977-221-67	Sequence 67, Appl
15	71	93.4	861	5 PCT-US95-06613-67	Sequence 67, Appl
16	69	90.8	873	3 US-09-187-331-6	Sequence 6, Appl
17	69	90.8	925	2 US-08-392-946-1	Sequence 1, Appl
18	69	90.8	925	2 US-08-504-169-1	Sequence 1, Appl
19	69	90.8	925	5 PCT-US94-14893-1	Sequence 1, Appl
20	40	52.6	203	2 US-08-684-024-8	Sequence 8, Appl
21	40	52.6	203	3 US-09-145-868-8	Sequence 8, Appl
22	40	52.6	205	2 US-08-684-024-1	Sequence 1, Appl
23	40	52.6	205	2 US-08-684-024-6	Sequence 6, Appl
24	40	52.6	205	2 US-08-684-024-7	Sequence 7, Appl
25	40	52.6	205	3 US-09-145-868-1	Sequence 1, Appl
26	40	52.6	205	3 US-09-145-868-6	Sequence 6, Appl
27	40	52.6	205	3 US-09-145-868-7	Sequence 7, Appl

28	38	50.0	438	3 US-09-187-331-2	Sequence 2, Appl
29	37	48.7	49	1 US-07-865-166A-6	Sequence 6, Appl
30	37	48.7	254	2 US-08-207-481-20	Sequence 20, Appl
31	37	48.7	254	5 PCT-US95-02689-20	Sequence 20, Appl
32	37	48.7	264	2 US-08-484-905-120	Sequence 120, App
33	37	48.7	264	3 US-08-481-985B-120	Sequence 120, App
34	37	48.7	264	4 US-08-370-476-120	Sequence 120, App
35	37	48.7	436	1 US-08-080-255-8	Sequence 8, Appl
36	37	48.7	436	4 US-08-465-713-8	Sequence 8, Appl
37	37	48.7	436	5 PCT-US93-05857-8	Sequence 8, Appl
38	37	48.7	3969	4 US-08-061-376-5	Sequence 5, Appl
39	36	47.4	309	3 US-08-969-644-20	Sequence 20, Appl
40	36	47.4	309	4 US-08-444-189-20	Sequence 20, Appl
41	36	47.4	359	2 US-08-560-398-10	Sequence 10, Appl
42	36	47.4	516	4 US-09-201-641-6	Sequence 6, Appl
43	36	47.4	707	1 US-08-528-122-18	Sequence 18, Appl
44	36	47.4	707	5 PCT-US95-11720-18	Sequence 18, Appl
45	36	47.4	750	1 US-08-325-553-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-346-455B-36  
; Sequence 36, Application US/08346455B  
; Patent No. 5731167  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA; DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,455B  
; FILING DATE: 28-NOV-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 788  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein



```

; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE: teratocarcinoma
; CELL LINE: N-tera 2D1
; ORGANELLE:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: N-tera 2D1 putative
; OTHER INFORMATION: ATX protein sequence
;
US-08-346-455B-36

Query Match 100.0%; Score 76; DB 1; Length 788;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13
Db 126 YMRPVYPTKTFPN 138

RESULT 2
US-08-977-221-36
; Sequence 36, Application US/08977221
; Patent No. 6084069
; GENERAL INFORMATION:
; APPLICANT: UNITED STATES OF AMERICA; DEPT.
; APPLICANT: OF HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: MOTILITY STIMULATING
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,221
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,455
; FILING DATE: 28-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,182
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/822,043
; FILING DATE: 17-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4149US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 36:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 788
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: Unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: No
; ORIGINAL SOURCE:
; ORGANISM: Human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE: teratocarcinoma
; CELL LINE: N-tera 2D1
; ORGANELLE:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: N-tera 2D1 putative
; OTHER INFORMATION: ATX protein sequence
;
US-08-977-221-36

Query Match 100.0%; Score 76; DB 3; Length 788;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13
Db 126 YMRPVYPTKTFPN 138

RESULT 3
PCT-US95-06613-36
; Sequence 36, Application PC/TUS9506613
; GENERAL INFORMATION:
; APPLICANT: STRACKE, MARY; LIOTTA, LANCE;
; APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,
; APPLICANT: HENRY; MURATA, JUN
; TITLE OF INVENTION: MOTILITY STIMULATING
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06613
; FILING DATE: 24-MAY-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,455
; FILING DATE: 28-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,182
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/822,043
; FILING DATE: 17-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH

```

REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US2  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 788  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE: teratocarcinoma  
CELL LINE: N-tera 2D1  
ORGANELLE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: N-tera 2D1 putative  
OTHER INFORMATION: ATX protein sequence  
PCT-US95-06613-36

Query Match 100.0%; Score 76; DB 5; Length 788;  
Best Local Similarity 100.0%; Pred. No. 0.00011;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
Db 126 YMRPVYPTKTFPN 138

RESULT 4  
US-08-346-455B-34  
Sequence 34, Application US/08346455B  
Patent No. 5731167  
GENERAL INFORMATION:  
APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/346,455B  
FILING DATE: 28-NOV-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182

FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149PCT  
TELEPHONE: (212) 751-6849  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 829  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE: Melanoma  
CELL LINE: A2058  
ORGANELLE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: Putative protein  
OTHER INFORMATION: sequence of A2058 Autotaxin  
US-08-346-455B-34

Query Match 100.0%; Score 76; DB 1; Length 829;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
Db 115 YMRPVYPTKTFPN 127

RESULT 5  
US-08-977-221-34  
Sequence 34, Application US/08977221  
Patent No. 6084069  
GENERAL INFORMATION:  
APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/977,221  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA: 08/346,455  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 829  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: No  
ORIGINAL SOURCE:  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE: Melanoma  
CELL LINE: A2058  
ORGANELLE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD: Putative protein  
OTHER INFORMATION: sequence of A2058 Autotaxin  
US-08-977-221-34

Query Match 100.0%; Score 76; DB 3; Length 829;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||||

Db 115 YMRPVYPTKTFPN 127

RESULT 6  
PCT-US95-06613-34  
Sequence 34, Application PC/TUS9506613  
GENERAL INFORMATION:  
APPLICANT: STRACKE, MARY; LIOTTA, LANCE;  
APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,  
APPLICANT: HENRY; MURATA, JUN  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/346,455  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 829  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: No  
ORIGINAL SOURCE:  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE: Melanoma  
CELL LINE: A2058  
ORGANELLE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: Putative protein  
OTHER INFORMATION: sequence of A2058 Autotaxin  
PCT-US95-06613-34

Query Match 100.0%; Score 76; DB 5; Length 829;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||||

Db 115 YMRPVYPTKTFPN 127

RESULT 7  
US-08-346-455B-69  
Sequence 69, Application US/08346455B  
Patent No. 5731167  
GENERAL INFORMATION:  
APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE



; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,455B  
; FILING DATE: 28-NOV-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 915  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: A2058 ATX protein  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION:  
US-08-346-455B-69

Query Match 100.0%; Score 76; DB 1; Length 915;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||||  
Db 201 YMRPVYPTKTFPN 213  
RESULT 8  
US-08-977-221-69  
; Sequence 69, Application US/08977221  
; Patent No. 6084069  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA; DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/977,221  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/346,455  
; FILING DATE: 28-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149US3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 915  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: A2058 ATX protein  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION:  
US-08-977-221-69

Query Match 100.0%; Score 76; DB 3; Length 915;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||||  
Db 201 YMRPVYPTKTFPN 213  
RESULT 9  
PCT-US95-06613-69  
; Sequence 69, Application PC/TUS9506613  
; GENERAL INFORMATION:  
; APPLICANT: STRACKE, MARY; LIOTTA, LANCE;  
; APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,  
; APPLICANT: HENRY; MURATA, JUN  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/346,455  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 915  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: No  
FEATURE:  
NAME/KEY: A2058 ATX protein  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PCT-US95-06613-69

Query Match 100.0%; Score 76; DB 5; Length 915;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
Db 201 YMRPVYPTKTFPN 213

RESULT 10  
US-08-346-455B-38  
Sequence 38, Application US/08346455B  
Patent No. 5731167  
GENERAL INFORMATION:  
APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/346,455B  
FILING DATE: 28-NOV-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613

FILING DATE: 24-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 979  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: No  
ORIGINAL SOURCE:  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: Liver  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: putative autotaxin  
OTHER INFORMATION: protein sequence from human liver  
US-08-346-455B-38

Query Match 100.0%; Score 76; DB 1; Length 979;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
Db 201 YMRPVYPTKTFPN 213

RESULT 11  
US-08-977-221-38  
Sequence 38, Application US/08977221  
Patent No. 6084069  
GENERAL INFORMATION:  
APPLICANT: UNITED STATES OF AMERICA; DEPT.  
APPLICANT: OF HEALTH AND HUMAN SERVICES  
TITLE OF INVENTION: MOTILITY STIMULATING  
TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1



; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,455B  
; FILING DATE: 28-NOV-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06613  
; FILING DATE: 24-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 751-6800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 861  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: N-tera 2D1 ATX protein  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION:  
; US-08-346-455B-67

Query Match 93.4%; Score 71; DB 1; Length 861;  
Best Local Similarity 92.3%; Pred. No. 0.00079;  
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 YMRPVYPTKTFPN 13  
Db 201 HMRPVYPTKTFPN 213  
  
RESULT 14  
US-08-977-221-67  
; Sequence 67, Application US/08977221  
; Patent No. 6084069  
; GENERAL INFORMATION:  
; APPLICANT: UNITED STATES OF AMERICA; DEPT.  
; APPLICANT: OF HEALTH AND HUMAN SERVICES  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK

; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/977,221  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/346,455  
; FILING DATE: 28-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249,182  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/822,043  
; FILING DATE: 17-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOROTHY R. AUTH  
; REGISTRATION NUMBER: 36,434  
; REFERENCE/DOCKET NUMBER: 2026-4149US3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 751-6800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 861  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: N-tera 2D1 ATX protein  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION:  
; US-08-977-221-67

Query Match 93.4%; Score 71; DB 3; Length 861;  
Best Local Similarity 92.3%; Pred. No. 0.00079;  
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
Db 201 HMRPVYPTKTFPN 213  
  
RESULT 15  
PCT-US95-06613-67  
; Sequence 67, Application PC/TUS9506613  
; GENERAL INFORMATION:  
; APPLICANT: STRACKE, MARY; LIOTTA, LANCE;  
; APPLICANT: SCHIFFMANN, ELLIOTT; KRUTZSCH,  
; APPLICANT: HENRY; MURATA, JUN  
; TITLE OF INVENTION: MOTILITY STIMULATING  
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND  
; TITLE OF INVENTION: THERAPY  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06613  
FILING DATE: 24-MAY-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/346,455  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/249,182  
FILING DATE: 25-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/822,043  
FILING DATE: 17-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4149US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 861  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: Unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: N-tera 2D1 ATX protein  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PCT-US95-06613-67

Query Match 93.4%; Score 71; DB 5; Length 861;  
Best Local Similarity 92.3%; Pred. No. 0.00079;  
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
:|||||  
Db 201 HMRPVYPTKTFPN 213

Search completed: July 23, 2001, 13:22:06  
Job time: 359 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2001, 14:45:54 ; Search time 25.97 seconds  
(without alignments)  
38.131 Million cell updates/sec

**Title:** US-09-483-831-69\_COPY\_201\_213

Perfect score: 76  
Sequence: 1 YMRPVVPTKTFPN 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database :      PIR_68:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	76	100.0	915	1	A55144	autotaxin precursor
2	70	92.1	885	1	A55453	plasma cell membra
3	69	90.8	96	2	A25274	phosphodiesterase
4	69	90.8	300	2	A41179	protein kinase PC-
5	69	90.8	925	1	A39216	plasma cell membra
6	66	86.8	905	1	A27410	plasma cell membra
7	65	85.5	875	1	A57080	cell surface antig
8	54	71.1	493	2	S50443	hypothetical prote
9	50	65.8	743	2	S19437	hypothetical prote
10	49	64.5	457	2	T09932	hypothetical prote
11	49	64.5	461	2	T09933	nucleotide pyropho
12	49	64.5	829	2	T19494	hypothetical prote
13	48	63.2	496	2	T09931	nucleotide pyropho
14	47	61.8	479	2	T03293	nucleotide pyropho
15	45	59.2	674	2	T19495	hypothetical prote
16	44.5	58.6	316	2	S46237	glucan endo-1,3-be
17	42	55.3	429	2	T33724	hypothetical prote
18	41	53.9	247	2	T32514	hypothetical prote
19	41	53.9	381	2	T23250	hypothetical prote
20	41	53.9	427	2	S57776	cysteine proteinas
21	41	53.9	433	2	B82537	phosphodiesterase
22	41	53.9	493	2	T01206	cysteine proteinas
23	41	53.9	1829	2	T24583	hypothetical prote
24	40	52.6	174	2	S15391	crustacyanin chain
25	40	52.6	205	2	G01942	mitotic feedback c
26	40	52.6	247	2	T45847	hypothetical prote
27	40	52.6	346	2	JA0159	cysteine proteinas
28	40	52.6	466	2	T06416	cysteine proteinas
29	40	52.6	1547	2	JQ0096	hypothetical 176K

30	39	51.3	269	2	A84841	probable embryo-ab
31	39	51.3	321	2	E86423	hypothetical prote
32	39	51.3	453	2	T16795	hypothetical prote
33	39	51.3	471	2	G69809	aminoacid carrier
34	39	51.3	485	2	T40657	probable phosphodie
35	39	51.3	528	2	F64580	hypothetical prote
36	39	51.3	528	2	F71931	outer membrane pro
37	39	51.3	1064	2	S52687	serine/threonine-s
38	38	50.0	256	2	T49254	Myb DNA binding pr
39	38	50.0	453	2	S67089	hypothetical prote
40	38	50.0	567	2	C72698	hypothetical prote
41	38	50.0	604	2	S24760	hydroxymethylgluta
42	38	50.0	708	2	T48022	hypothetical prote
43	38	50.0	791	2	T39924	hypothetical prote
44	38	50.0	1068	2	E81965	probable outer mem
45	38	50.0	1082	2	H81020	serotype-1-specific

## ALIGNMENTS

RESULT 1  
A55144  
autotaxin precursor - human  
N:Contains: phosphodiesterase I (EC 3.1.4.1)  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A55144; A42329  
R:Murata, J.; Lee, H.Y.; Clair, T.; Krutzsch, H.C.; Arestad, A.A.; Sobel, M.E.; Liott  
J. Biol. Chem. 269, 30479-30484, 1994  
A:Title: cDNA cloning of the human tumor motility-stimulating protein, autotaxin, rev  
A:Reference number: A55144; MUID:95074054  
A:Accession: A55144  
A:Molecule type: mRNA  
A:Residues: 1-915 <MUR>  
A:Cross-references: GB:L35594; NID:g537905; PIDN:AAA64785.1; PID:g537906  
A:Note: parts of this sequence were confirmed by peptide sequencing  
R:Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.; Cioce, V.; Schiffmann,  
J. Biol. Chem. 267, 2524-2529, 1992  
A:Title: Identification, purification, and partial sequence analysis of autotaxin, a  
A:Reference number: A42329; MUID:92129337  
A:Accession: A42329  
A:Molecule type: protein  
A:Residues: 256-266;422-444;504-507,'AN';510,'X',511-515;533-548;'S',554-559,'N',561-  
A:Experimental source: A2058 melanoma cells  
A:Note: sequence extracted from NCBI backbone (NCBIP:78526, NCBIP:78523, NCBIP:78521,  
A:Note: a peptide fragment Tyr-Asp-Val-Pro-Trp-Asn-Glu-Thr-Ile was also found  
C:Comment: This protein acts as an autocrine factor to stimulate tumor cell motility.  
C:Genetics:  
A:Gene: GDB:ATX  
A:Cross-references: GDB:378346  
A:Map position: 8q22-8qter  
C:Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C:Keywords: EF hand; glycoprotein; phosphoprotein; phosphoric diester hydrolase  
F:55-98/Domain: somatomedin B homology <SBH1>  
F:99-142/Domain: somatomedin B homology <SBH2>  
F:54,463,577,859/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:210/Binding site: AMP (Thr) (covalent) #status predicted

```
Query Match      100.0%; Score 76; DB 1; Length 915;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 YMRPVYPTKTEPN 13  
 |||||  
 Db 201 YMRPVYPTKTEPN 213

RESULT 2  
A55453  
plasma cell membrane glycoprotein PC-1, brain specific - rat  
N:Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)

C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C;Accession: A55453; JU0187  
R;Narita, M.; Goji, J.; Nakamura, H.; Sano, K.  
J. Biol. Chem. 269, 28235-28242, 1994  
A;Title: Molecular cloning, expression, and localization of a brain-specific phosphodies  
A;Reference number: A55453; MUID:95050605  
A;Accession: A55453  
A;Molecule type: mRNA  
A;Residues: 1-885 <NAR>  
A;Cross-references: GB:D28560; NID:g464196; PIDN:BAA05910.1; PID:g464197  
R;Narita, M.; Goji, J.; Sano, K.; Nakamura, H.  
submitted to JIPID, February 1994  
A;Description: Cloning and expression of brain-specific nucleotide diphosphohydrolase.  
A;Reference number: JU0187  
A;Accession: JU0187  
A;Molecule type: mRNA  
A;Residues: 1-66, 'Q', 68-81, 'T', 83-94, 'C', 96, 'A', 98-195, 'A', 197-514, 'E', 516-621, 'E', 623-6  
A;Experimental source: strain Sprague-Dawley  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: coenzyme A; exonuclease; glycoprotein; phosphoprotein; phosphoric diester hy  
F;54-97/Domain: somatomedin B homology <SBH1>  
F;98-141/Domain: somatomedin B homology <SBH2>  
F;53,150,396,408,522,608,829/Binding site: carbohydrate (Asn) (covalent) #status predict  
F;207/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 92.1%; Score 70; DB 1; Length 885;  
Best Local Similarity 92.3%; Pred. No. 0.00058;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | |  
Db 198 YTRPVYPTKTFPN 210

RESULT 3  
A25274  
phosphodiesterase I (EC 3.1.4.1) - bovine (fragments)  
N;Alternate names: 5'-exonuclease  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 21-May-1988 #sequence\_revision 23-Mar-1995 #text\_change 21-May-1999  
C;Accession: A25274; C48395; A48395; D48395  
R;Culp, J.S.; Blytt, H.J.; Hermodson, M.; Butler, L.G.  
J. Biol. Chem. 260, 8320-8324, 1985  
A;Title: Amino acid sequence of the active site peptide of bovine intestinal 5'-nucleoti  
A;Reference number: A25274; MUID:85234541  
A;Accession: A25274  
A;Molecule type: protein  
A;Residues: 1-61 <CUL>  
R;Maruyama, E.; Iwamatsu, A.; Takashima, S.  
Biochem. Mol. Biol. Int. 29, 579-586, 1993  
A;Title: Purification and amino acid microsequencing of alkaline phosphodiesterase I fro  
A;Reference number: A48395; MUID:93250579  
A;Accession: C48395  
A;Molecule type: protein  
A;Residues: 22-36;62-73;74-81;82-96 <MAR>  
A;Experimental source: kidney  
A;Note: sequence extracted from NCBI backbone (NCBIP:132415, NCBIP:131478)  
A;Note: the correct order of the fragments is unknown  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: phosphoprotein; phosphoric diester hydrolase  
F;39/Binding site: AMP (Thr) (covalent) #status experimental

Query Match 90.8%; Score 69; DB 2; Length 96;  
Best Local Similarity 84.6%; Pred. No. 9.1e-05;  
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | |  
Db 30 YLRPAYPTKTFPN 42

RESULT 4  
A41179

protein kinase PC-1 (EC 2.7.1.-) - bovine (fragments)  
N;Alternate names: MAFF; major acidic fibroblast growth factor-stimulated phosphoprot  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 28-May-1992 #sequence\_revision 22-Apr-1995 #text\_change 30-Apr-1999  
C;Accession: A41179; A49308  
R;Oda, Y.; Kuo, M.D.; Huang, S.S.; Huang, J.S.  
J. Biol. Chem. 266, 16791-16795, 1991  
A;Title: The plasma cell membrane glycoprotein, PC-1, is a threonine-specific protein  
A;Reference number: A41179; MUID:91358477  
A;Accession: A41179  
A;Molecule type: protein  
A;Residues: 1-26;36-56;59-67;68-133;134-144;145-267;268-300 <ODA>  
A;Experimental source: liver  
R;Oda, Y.; Kuo, M.D.; Huang, S.S.; Huang, J.S.  
J. Biol. Chem. 268, 27318-27326, 1993  
A;Title: The major acidic fibroblast growth factor (aFGF)-stimulated phosphoprotein f  
estrase activities.  
A;Reference number: A49308; MUID:94086550  
A;Accession: A49308  
A;Molecule type: protein  
A;Residues: 27-35, 'X', 37-58 <OD2>  
A;Experimental source: liver  
A;Note: sequence extracted from NCBI backbone (NCBIP:141583)  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: glycoprotein; phosphoprotein; phosphotransferase  
F;1-25/Domain: somatomedin B homology (fragment) <SBH>  
F;36/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 90.8%; Score 69; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MRPVYPTKTFPN 13  
| | | | | | | | | |  
Db 28 MRPVYPTKTFPN 39

RESULT 5  
A39216

plasma cell membrane glycoprotein PC-1 - human  
N;Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)  
C;Species: Homo sapiens (man)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-May-2000  
C;Accession: A39216; S21706; S23587; S51030  
R;Buckley, M.F.; Loveland, K.A.; McKinstry, W.J.; Garson, O.M.; Goding, J.W.  
J. Biol. Chem. 265, 17506-17511, 1990  
A;Title: Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human molecule,  
A;Reference number: A39216; MUID:91009202  
A;Accession: A39216  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-925 <BUC>  
A;Cross-references: GB:J05654  
R;Funakoshi, I.; Kato, H.; Horie, K.; Yano, T.; Hori, Y.; Kobayashi, H.; Inoue, T.; S  
Arch. Biochem. Biophys. 295, 180-187, 1992  
A;Title: Molecular cloning of cDNAs for human fibroblast nucleotide pyrophosphatase.  
A;Reference number: S21706; MUID:92246539  
A;Accession: S21706  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-925 <FUN1>  
A;Accession: S23587  
A;Molecule type: protein  
A;Residues: 116-121;247-271, 'X', 273-275;279-280, 'X', 282-283;303-316;362-364;449-465;4

A;Note: it is uncertain whether Met-1 or Met-53 is the initiator  
R;Belli, S.I.; Goding, J.W.  
Eur. J. Biochem. 226, 433-443, 1994  
A;Title: Biochemical characterization of human PC-1, an enzyme possessing alkaline ph

A;Reference number: S51030; MUID:95094801  
A;Accession: S51030  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-80 <BEL>  
C;Genetics:

A;Gene: GDB:PDNPL1; M6S1; NPPS  
A;Cross-references: GDB:132615; OMIM:173335  
A;Map position: 6q22-6q23  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: coenzyme A; glycoprotein; phosphoprotein; phosphoric diester hydrolase; transmembrane #status predicted <TMM>  
F;77-97/Domain: transmembrane #status predicted <TMM>  
F;104-144/Domain: somatomedin B homology <SBH1>  
F;145-188/Domain: somatomedin B homology <SBH2>  
F;179,285,341,477,578,585,643,700,731,748/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;256/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 90.8%; Score 69; DB 1; Length 925;  
Best Local Similarity 100.0%; Pred. No. 0.0009;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MRPVYPTKTFPN 13  
|||||  
Db 248 MRPVYPTKTFPN 259

RESULT 6  
A27410

plasma cell membrane glycoprotein PC-1 - mouse  
N;Alternate names: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.1.3)  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-May-2000  
C;Accession: A27410; I59055; S38354  
R;van Driel, I.R.; Goding, J.W.  
J. Biol. Chem. 262, 4882-4887, 1987  
A;Title: Plasma cell membrane glycoprotein PC-1. Primary structure deduced from cDNA cloning  
A;Reference number: A27410; MUID:87165906

A;Accession: A27410  
A;Molecule type: mRNA  
A;Residues: 1-905 <VAN>  
A;Cross-references: GB:J02700; NID:g200236; PIDN:AAA39893.1; PID:g200237  
A;Note: the authors translated the codon CAG for residue 24 as Glu  
R;van Driel, I.R.; Wilks, A.F.; Pietersz, G.A.; Goding, J.W.  
Proc. Natl. Acad. Sci. U.S.A. 82, 8619-8623, 1985  
A;Title: Murine plasma cell membrane antigen PC-1: Molecular cloning of cDNA and analysis  
A;Reference number: I59055; MUID:86094275  
A;Accession: I59055  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 203-219 <RES>  
A;Cross-references: GB:M12552; NID:g200234; PIDN:AAA39892.1; PID:g200235  
R;Belli, S.I.; van Driel, I.R.; Goding, J.W.  
Eur. J. Biochem. 217, 421-428, 1993

A;Title: Identification and characterization of a soluble form of the plasma cell membrane antigen PC-1  
A;Reference number: S38354; MUID:94039066  
A;Accession: S38354  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 35-219 <BEL>  
A;Cross-references: EMBL:L04516  
C;Genetics:  
A;Introns: 62/3; 87/1; 126/1; 168/1; 188/2  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: coenzyme A; glycoprotein; phosphoprotein; phosphoric diester hydrolase; transmembrane #status predicted <SBH1>  
F;86-126/Domain: somatomedin B homology <SBH1>  
F;127-170/Domain: somatomedin B homology <SBH2>  
F;161,267,323,459,567,624/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;238/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 86.8%; Score 66; DB 1; Length 905;  
Best Local Similarity 91.7%; Pred. No. 0.0028;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 MRPVYPTKTFPN 13  
|||||  
Db 230 MRPVYPTKTFPN 241

RESULT 7  
A57080

cell surface antigen RB13-6 - rat  
N;Contains: phosphodiesterase I (EC 3.1.4.1)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A57080  
R;Deissler, H.; Lottspeich, F.; Rajewsky, M.F.  
J. Biol. Chem. 270, 9849-9855, 1995  
A;Title: Affinity purification and cDNA cloning of rat neural differentiation and tumorigenesis  
A;Reference number: A57080; MUID:95247775  
A;Accession: A57080  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-875 <DEI>

A;Cross-references: GB:Z47987; NID:g806378; PIDN:CAA88029.1; PID:g806379  
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology  
C;Keywords: exonuclease; glycoprotein; phosphoprotein; phosphoric diester hydrolase; transmembrane #status predicted <CYT>  
F;1-22/Domain: cytosolic #status predicted <CYT>  
F;23-45/Domain: transmembrane #status predicted <TMM>  
F;46-875/Domain: extracellular #status predicted <EXT>  
F;51-94/Domain: somatomedin B homology <SBH>  
F;95-138/Domain: somatomedin B homology <SBH2>  
F;206/Binding site: AMP (Thr) (covalent) #status predicted  
F;237,280,289,533,574,594,702,789/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.5%; Score 65; DB 1; Length 875;  
Best Local Similarity 84.6%; Pred. No. 0.0041;  
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||||  
Db 197 YMRPVYPTKTFPN 209

RESULT 8  
S50443

hypothetical protein YEL016c - yeast (Saccharomyces cerevisiae)  
C;Species: Saccharomyces cerevisiae  
C;Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 29-Oct-1999  
C;Accession: S50443  
R;Dietrich, F.S.  
submitted to the EMBL Data Library, December 1994  
A;Description: Saccharomyces cerevisiae chromosome V cosmids 9871, 8199, 9867, 9495 a  
A;Reference number: S50428  
A;Accession: S50443  
A;Molecule type: DNA  
A;Residues: 1-493 <DIE>  
A;Cross-references: EMBL:U18530; NID:g602367; PID:g602383; GSPDB:GN00005; MIPS:YEL016  
C;Genetics:  
A;Gene: MIPS:YEL016c  
A;Map position: 5L

Query Match 71.1%; Score 54; DB 2; Length 493;  
Best Local Similarity 69.2%; Pred. No. 0.17;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||||  
Db 118 YMIPSFTQTTFPN 130

RESULT 9  
S19437



hypothetical protein YCR026c - yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein YCR247  
C;Species: Saccharomyces cerevisiae  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 12-Dec-1997  
C;Accession: S19437; S19750; S27380  
R;Pohl, F.; Richterich, P.; Wurst, H.  
submitted to the Protein Sequence Database, March 1992  
A;Reference number: S19437  
A;Accession: S19437  
A;Molecule type: DNA  
A;Residues: 1-244 <POH>  
A;Cross-references: EMBL:X59720; MIPS:YCR026c  
R;Berben, G.; Bolle, P.A.; Gilliquet, V.; Hilger, F.  
submitted to the Protein Sequence Database, March 1992  
A;Reference number: S19433  
A;Accession: S19750  
A;Molecule type: DNA  
A;Residues: 244-743 <BER>  
A;Cross-references: EMBL:X59720; MIPS:YCR026c  
R;Bolle, P.A.; Gilliquet, V.; Berben, G.; Dumont, J.; Hilger, F.  
Yeast 8, 205-213, 1992  
A;Title: The complete sequence of K3B, a 7.9 kb fragment between PGK1 and CRY1 on chromosome 8  
A;Reference number: S22273; MUID:92245758  
A;Accession: S27380  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-106,'H',108-119,'F',121-211,'V',213-743 <BOL>  
C;Genetics:  
A;Map position: 3R  
C;Keywords: transmembrane protein

Query Match 65.8%; Score 50; DB 2; Length 743;  
Best Local Similarity 61.5%; Pred. No. 1.2;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
:|:|:|:|:|  
Db 210 FMIPSPFTTFPN 222

RESULT 10  
T09932  
nucleotide pyrophosphatase homolog T16L4.200 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999  
C;Accession: T09932  
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z16897  
A;Accession: T09932  
A;Molecule type: DNA  
A;Residues: 1-457 <BEV>  
A;Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.200  
A;Experimental source: cultivar Columbia; BAC clone T16L4  
C;Genetics:  
A;Gene: ATSP:T16L4.200  
A;Map position: 4

Query Match 64.5%; Score 49; DB 2; Length 457;  
Best Local Similarity 80.0%; Pred. No. 1.1;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PVYPTKTFPN 13  
|:|:|:|:|  
Db 94 PVFPTMTFPN 103

RESULT 11  
T09933  
nucleotide pyrophosphatase homolog T16L4.210 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999  
C;Accession: T09933  
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z16897  
A;Accession: T09933  
A;Molecule type: DNA  
A;Residues: 1-461 <BEV>  
A;Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.210  
A;Experimental source: cultivar Columbia; BAC clone T16L4  
C;Genetics:  
A;Gene: ATSP:T16L4.210  
A;Map position: 4

Query Match 64.5%; Score 49; DB 2; Length 461;  
Best Local Similarity 80.0%; Pred. No. 1.1;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PVYPTKTFPN 13  
|:|:|:|:|  
Db 97 PVFPTMTFPN 106

RESULT 12  
T19494  
hypothetical protein C27A7.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T19494  
R;Harris, B.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: Z19132  
A;Accession: T19494  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-829 <WIL>  
A;Cross-references: EMBL:Z81041; PIDN:CAB02784.1; GSPDB:GN00023; CESP:C27A7.1  
A;Experimental source: clone C27A7  
C;Genetics:  
A;Gene: CESP:C27A7.1  
A;Map position: 5  
A;Introns: 12/3; 37/1; 100/3; 176/2; 272/3; 392/1; 427/1; 479/2; 566/2; 588/3; 677/1;

Query Match 64.5%; Score 49; DB 2; Length 829;  
Best Local Similarity 80.0%; Pred. No. 2;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PVYPTKTFPN 13  
|:|:|:|:|  
Db 218 PSYPSKTFPN 227

RESULT 13  
T09931  
nucleotide pyrophosphatase homolog T16L4.190 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999  
C;Accession: T09931  
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z16897  
A;Accession: T09931  
A;Molecule type: DNA  
A;Residues: 1-496 <BEV>  
A;Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.190  
A;Experimental source: cultivar Columbia; BAC clone T16L4  
C;Genetics:  
A;Gene: ATSP:T16L4.190  
A;Map position: 4

Query Match 63.2%; Score 48; DB 2; Length 496;  
Best Local Similarity 80.0%; Pred. NO. 1.8;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PVYPTKTFPN 13  
|:|:|:|:|  
Db 146 PVFPTLTFPN 155

RESULT 14  
T03293  
nucleotide pyrophosphatase homolog - rice  
C:Species: Oryza sativa (rice)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C:Accession: T03293  
R:Hsing, Y.C.; Tsao, C.V.; Chow, T.; Hsieh, J.; Chen, Z.  
submitted to the EMBL Data Library, April 1995  
A:Description: Rice early embryogenesis gene.  
A:Reference number: Z14889  
A:Accession: T03293  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-479 <HSI>  
A:Cross-references: EMBL:U25430; NID:g818848; PID:g818849  
A:Experimental source: strain Tainung 67  
C:Genetics:  
A:Note: OSE4

Query Match 61.8%; Score 47; DB 2; Length 479;  
Best Local Similarity 70.0%; Pred. No. 2.5;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PVYPTKTFPN 13  
|:|:|:|:|  
Db 129 PIFPTLTFPN 138

RESULT 15  
T19495  
hypothetical protein C27A7.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T19495  
R:Harris, B.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19132  
A:Accession: T19495  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-674 <WIL>  
A:Cross-references: EMBL:Z81041; PIDN:CAB02785.1; GSPDB:GN00023; CESP:C27A7.3  
A:Experimental source: clone C27A7  
C:Genetics:  
A:Gene: CESP:C27A7.3  
A:Map position: 5  
A:Introns: 61/1; 98/3; 191/1; 220/3; 268/3; 340/1; 392/2; 475/2; 496/3; 537/1; 588/1

Query Match 59.2%; Score 45; DB 2; Length 674;  
Best Local Similarity 70.0%; Pred. NO. 7.7;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PVYPTKTFPN 13  
|:|:|:|:|  
Db 166 PSFPSKTFPN 175



OX NCB1\_TaxID=9606;  
RN [1]  
RX SEQUENCE FROM N.A.  
RA MEDLINE=91009202; PubMed=2211644;  
RA Buckley M.F., Loveland K.A., McKinstry W.J., Garson O.M., Goding J.W.;  
RT "Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human  
molecule, amino acid sequence, and chromosomal location.";  
RL J. Biol. Chem. 265:17506-17511(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92246539; PubMed=1315502;  
RA Funakoshi I., Kato H., Horie K., Yano T., Hori Y., Kobayashi H.,  
RA Inoue T., Suzuki H., Fukui S., Tsukahara M., Kajii T.,  
RA Yamashina I.;  
RT "Molecular cloning of cDNAs for human fibroblast nucleotide  
pyrophosphatase.";  
RL Arch. Biochem. Biophys. 295:180-187(1992).  
RN [3]  
RP SEQUENCE OF 93-133 FROM N.A., AND VARIANT GLN-121.  
RX MEDLINE=99408501; PubMed=10480624;  
RA Pizzuti A., Frittitta L., Argiolas A., Baratta R., Goldfine I.D.,  
RA Bozzali M., Ercolino T., Scariato G., Iacoviello L., Vigneri R.,  
RA Tassi V., Trischitta V.;  
RT "A polymorphism (K121Q) of the human glycoprotein PC-1 gene coding  
region is strongly associated with insulin resistance.";  
RL Diabetes 48:1881-1884(1999).  
CC -!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.  
CC -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES  
CC SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED  
CC OLIGO-NUCLEOTIDES.  
CC -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER  
CC OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE  
CC OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.  
CC -!- SIMILARITY: CONTAINS 2 SOMATOMEDIN-B TYPE DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M57736; AAA63237.1; -.  
DR EMBL; D12485; BAA02054.1; -.  
DR EMBL; D12485; BAA02053.1; ALT\_INIT.  
DR EMBL; AF067177; AAD38420.1; -.  
DR EMBL; AF067178; AAD38421.1; -.  
DR PIR; A39216; A39216.  
DR MIM; 173335; -.  
DR InterPro; IPR001212; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01663; Phosphodiesterase; 1.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR PRINTS; PR00022; SOMATOMEDINB.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
KW Glycoprotein; Transmembrane; Duplication; Signal-anchor; Hydrolase;  
KW Polymorphism.  
FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 25 45 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
FT DOMAIN 46 873 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 52 92 SOMATOMEDIN-B LIKE.  
FT DOMAIN 93 136 SOMATOMEDIN-B LIKE.  
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 533 533 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 696 696 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 121 121 K -> Q.  
FT /FTid=VAR\_008873.  
SQ SEQUENCE 873 AA; 99929 MW; 872808C20B048070 CRC64;  
  
Query Match 90.8%; Score 69; DB 1; Length 873;  
Best Local Similarity 100.0%; Pred. No. 0.00031;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 MRPVYPTKTFPN 13  
|||||  
Db 196 MRPVYPTKTFPN 207  
  
RESULT 3  
PC1\_MOUSE  
ID PC1\_MOUSE STANDARD; PRT; 871 AA.  
AC P06802;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 [INCLUDES: ALKALINE  
DE PHOSPHODIESTERASE I (EC 3.1.4.1); NUCLEOTIDE PYROPHOSPHATASE  
DE (EC 3.6.1.9) (NPPASE)].  
GN PDNPI OR PC1 OR NPPS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCB1\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87165906; PubMed=3104326;  
RA van Driel I.R., Goding J.W.;  
RT "Plasma cell membrane glycoprotein PC-1. Primary structure deduced  
RT from cDNA clones.";  
RL J. Biol. Chem. 262:4882-4887(1987).  
RN [2]  
RP PARTIAL SEQUENCE.  
RX MEDLINE=85056299; PubMed=3917281;  
RA Stearne P.A., van Driel I.R., Grego B., Simpson R.J., Goding J.W.;  
RT "The murine plasma cell antigen PC-1: purification and partial amino  
RT acid sequence.";  
RL J. Immunol. 134:443-448(1985).  
RN [3]  
RP FUNCTION, AND SEQUENCE FROM N.A.  
RX MEDLINE=91271356; PubMed=1647027;  
RA Rebbe N.F., Tong B.D., Finley E.M., Hickman S.;  
RT "Identification of nucleotide pyrophosphatase/alkaline  
RT phosphodiesterase I activity associated with the mouse plasma cell  
RT differentiation antigen PC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5192-5196(1991).  
RN [4]  
RP SEQUENCE OF 169-185 FROM N.A.  
RX MEDLINE=86094275; PubMed=3001713;  
RA van Driel I.R., Wilks A.F., Pietersz G.A., Goding J.W.;  
RT "Murine plasma cell membrane antigen PC-1: molecular cloning of cDNA  
RT and analysis of expression.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:8619-8623(1985).  
CC -!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.  
CC -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES  
CC SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED  
CC OLIGO-NUCLEOTIDES.  
CC -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.  
CC -!- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED ON THE SURFACE OF  
CC ANTIBODY-SECRETING CELLS.  
CC -!- SIMILARITY: CONTAINS 2 SOMATOMEDIN-B TYPE DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; J02700; AAA39893.1; ALT\_INIT.  
DR EMBL; M12552; AAA39892.1; -.  
DR PIR; A27410; A27410.  
DR MGD; MGI:97370; Pdnpl.  
DR InterPro; IPR001212; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01663; Phosphodiester; 1.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR PRINTS; PR00022; SOMATOMEDINB.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
KW Glycoprotein; Transmembrane; Duplication; Signal-anchor; Hydrolase.  
FT MOD\_RES ?1 ?1  
FT DOMAIN 1 24  
FT TRANSMEM 25 45  
FT DOMAIN 46 871  
FT DOMAIN 52 92  
FT DOMAIN 93 136  
FT CARBOHYD 127 127  
FT CARBOHYD 233 233  
FT CARBOHYD 289 289  
FT CARBOHYD 425 425  
FT CARBOHYD 533 533  
FT CARBOHYD 590 590  
FT SEQUENCE 871 AA; 99487 MW; 80848F81071F70AA CRC64;

Query Match 86.8%; Score 66; DB 1; Length 871;  
Best Local Similarity 91.7%; Pred. No. 0.001;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MRPVYPTKTFPN 13  
|||:|||||  
Db 196 MRPMPYPTKTFPN 207

RESULT 4  
YEB6\_YEAST  
ID YEB6\_YEAST STANDARD; PRT; 493 AA.  
AC P39997;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE HYPOTHEICAL 57.4 KDA PROTEIN IN PMP2-VAC8 INTERGENIC REGION.  
GN YEL016C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,  
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: TO YEAST YCR26C.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC -----

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U18530; AAB64493.1; -.  
DR SGD; S0000742; YEL016C.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01663; Phosphodiester; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 493 AA; 57354 MW; 10E67A05C6DEDF09 CRC64;

Query Match 71.1%; Score 54; DB 1; Length 493;  
Best Local Similarity 69.2%; Pred. No. 0.063;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||:|||||  
Db 118 YMIPSFPTQTTFPN 130

RESULT 5  
YCR6\_YEAST  
ID YCR6\_YEAST STANDARD; PRT; 743 AA.  
AC P25353;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE HYPOTHEICAL 84.9 KDA PROTEIN IN PMP1-FEN2 INTERGENIC REGION.  
GN YCR026C OR YCR26C OR YCR246.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE OF 1-244 FROM N.A.  
RA Pohl F., Richterich P., Wurst H.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 244-743 FROM N.A.  
RX MEDLINE=92245758; PubMed=1574926;  
RA Bolle P.-A., Gilliquet V., Berben G., Dumont J., Hilger F.;  
RT "The complete sequence of K3B, a 7.9 kb fragment between PGK1 and  
RT CRY1 on chromosome III, reveals the presence of seven open reading  
RT frames.";  
RL Yeast 8:205-213(1992).  
CC -!- SIMILARITY: TO YEAST YEL016C.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X59720; CAA42318.1; -.  
DR PIR; S19437; S19437.  
DR PIR; S27380; S27380.  
DR SGD; S0000621; YCR026C.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01663; Phosphodiester; 1.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 114 135  
FT SEQUENCE 743 AA; 84915 MW; 2385E9F4BA948B0D CRC64;  
POTENTIAL.

Query Match 65.8%; Score 50; DB 1; Length 743;  
Best Local Similarity 61.5%; Pred. No. 0.47;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||:|||||  
Db 210 FMIPSFPTTFPN 222

```
RESULT 6
E13E_HORVU
ID E13E_HORVU STANDARD; PRT; 316 AA.
AC Q02438;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GLUCAN ENDO-1,3-BETA-GLUCOSIDASE GV (EC 3.2.1.39) ((1->3)-BETA-GLUCAN
DE ENDOHYDROLASE GV) ((1->3)-BETA-GLUCANASE ISOENZYME GV) (BETA-1,3-
DE ENDOGLUCANASE GV).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae;
OC Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CLIPPER; TISSUE=Root, and Leaf;
RX MEDLINE=94307426; PubMed=8034043;
RA Xu P., Harvey A.J., Fincher G.B.;
RT "Heterologous expression of cDNAs encoding barley (Hordeum vulgare)
RT (1->3)-beta-glucanase isoenzyme GV.";
RL FEBS Lett. 348:206-210(1994).
RN [2]
RP SEQUENCE OF 5-316 FROM N.A.
RC STRAIN=CV. CLIPPER; TISSUE=Root, and Leaf;
RX MEDLINE=93013030; PubMed=1398132;
RA Xu P., Wang J., Fincher G.B.;
RT "Evolution and differential expression of the (1->3)-beta-glucan
RT endohydrolase-encoding gene family in barley, Hordeum vulgare.";
RL Gene 120:157-165(1992).
CC -!- FUNCTION: MAY PROVIDE A DEGREE OF PROTECTION AGAINST MICROBIAL
CC INVASION OF GERMINATED BARLEY GRAIN THROUGH ITS ABILITY TO DEGRADE
CC FUNGAL CELL WALL POLYSACCHARIDES.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,3-BETA-D-GLUCOSIDIC LINKAGES
CC IN 1,3-BETA-D-GLUCANS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -!- SIMILARITY: BELONGS TO FAMILY 17 OF GLYCOSYL HYDROLASES.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M96939; AAA21564.1; -.
PIR; JC1438; JC1438.
HSSP; P15737; LGHS.
InterPro; IPR00490; -.
Pfam; PF00332; Glyco_hydro_17; 1.
DR PROSITE; PS00587; GLYCOSYL_HYDROL_F17; 1.
KW Hydrolase; Glycosidase; Multigene family.
FT ACT_SITE 239 239 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 296 296 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 316 AA; 34413 MW; 0CC0AA9D48269B4F CRC64;

Query Match 58.6%; Score 44.5; DB 1; Length 316;
Best Local Similarity 81.8%; Pred. No. 1.7;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 3 RPVYPTKTFPN 13
:|||||
Db 307 QPVYPT-TFPN 316

RESULT 7
BGAL_THEET
ID BGAL_THEET STANDARD; PRT; 743 AA.
AC P77989;

Query Match 58.6%; Score 44.5; DB 1; Length 316;
Best Local Similarity 81.8%; Pred. No. 1.7;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 3 RPVYPTKTFPN 13
:|||||
Db 307 QPVYPT-TFPN 316

RESULT 8
CRA2_HOMGA
ID CRA2_HOMGA STANDARD; PRT; 174 AA.
AC P80007;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE CRUSTACYANIN A2 SUBUNIT.
OS Homarus gammarus (European lobster) (Homarus vulgaris).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6707;
RN [1]
RP SEQUENCE.
RX MEDLINE=91224133; PubMed=2026162;
RA Keen J.N., Caceres I., Eliopoulos E.E., Zagalsky P.F.,
RA Findlay J.B.C.;
RT "Complete sequence and model for the A2 subunit of the carotenoid
RT pigment complex, crustacyanin.";
RL Eur. J. Biochem. 197:407-417(1991).
CC -!- FUNCTION: BINDS THE CAROTENOID ASTAXANTHIN WHICH PROVIDES THE BLUE
CC COLORATION TO THE CARAPACE OF THE LOBSTER.
CC -!- SUBUNIT: OLIGOMER; CAN FORM DIMERS (BETA-CRUSTACYANIN); OR
CC COMPLEXES OF 16 SUBUNITS (ALPHA-CRUSTACYANIN). THERE ARE FIVE
CC TYPES OF SUBUNITS: A1, A2, A3, C1 AND C2.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR, CARAPACE.
```

```
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE BETA-GALACTOSIDASE (EC 3.2.1.23) (LACTASE).
GN LACZ OR LACA.
OS Thermoanaerobacter ethanolicus (Clostridium thermohydrosulfuricum).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Thermoanaerobacter group; Thermoanaerobacter.
OX NCBI_TaxID=1757;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33223 / 39E;
RA Zverlov V.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
CC GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Y08557; CAA69850.1; -.
InterPro; IPR001649; -.
Pfam; PF00703; Glyco_hydro_2; 1.
DR PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 388 388 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 453 453 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 743 AA; 85796 MW; FE01FF517E51DFC CRC64;

Query Match 55.3%; Score 42; DB 1; Length 743;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13
||:||||:|
Db 454 YMGHMYPTKSYDN 466

RESULT 8
CRA2_HOMGA
ID CRA2_HOMGA STANDARD; PRT; 174 AA.
AC P80007;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE CRUSTACYANIN A2 SUBUNIT.
OS Homarus gammarus (European lobster) (Homarus vulgaris).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6707;
RN [1]
RP SEQUENCE.
RX MEDLINE=91224133; PubMed=2026162;
RA Keen J.N., Caceres I., Eliopoulos E.E., Zagalsky P.F.,
RA Findlay J.B.C.;
RT "Complete sequence and model for the A2 subunit of the carotenoid
RT pigment complex, crustacyanin.";
RL Eur. J. Biochem. 197:407-417(1991).
CC -!- FUNCTION: BINDS THE CAROTENOID ASTAXANTHIN WHICH PROVIDES THE BLUE
CC COLORATION TO THE CARAPACE OF THE LOBSTER.
CC -!- SUBUNIT: OLIGOMER; CAN FORM DIMERS (BETA-CRUSTACYANIN); OR
CC COMPLEXES OF 16 SUBUNITS (ALPHA-CRUSTACYANIN). THERE ARE FIVE
CC TYPES OF SUBUNITS: A1, A2, A3, C1 AND C2.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR, CARAPACE.
```

```
CC -!- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
DR PIR; S15391; S15391.
DR InterPro; IPR000566; -.
DR InterPro; IPR002345; -.
DR InterPro; IPR003057; -.
DR Pfam; PF00061; lipocalin; 1.
DR PRINTS; PR00179; LIPOCALIN.
DR PRINTS; PR01273; INVTBRTCOLOR.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Pigment; Lipocalin; Transport.
FT DISULFID 12 119 BY SIMILARITY.
FT DISULFID 46 170 BY SIMILARITY.
SQ SEQUENCE 174 AA; 19670 MW; AC47FAA650C5E44E CRC64;

Query Match 52.6%; Score 40; DB 1; Length 174;
Best Local Similarity 87.5%; Pred. No. 5.2;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 VYPTKTFP 12
   |||||
Db 80 VYPTKEFP 87

RESULT 9
MD21_HUMAN
ID MD21_HUMAN STANDARD; PRT; 205 AA.
AC Q13257;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MITOTIC SPINDLE ASSEMBLY CHECKPOINT PROTEIN MAD2A (MAD2-LIKE 1)
DE (HSMAD2).
GN MAD2L1 OR MAD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=96421709; PubMed=8824189;
RA Li Y., Benezra R.;
RT "Identification of a human mitotic checkpoint gene: hSMAD2.";
RL Science 274:246-248(1996).
RN [2];
RP SEQUENCE FROM N.A.
RA Jin D.-Y., Jeang K.-T.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBBJ databases.
RN [3];
RP SEQUENCE FROM N.A.
RA Kleibert S., Barnikol-Watanabe S., Kratzin H.D., Hilschmann N.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBBJ databases.
RN [4];
RP INTERACTION WITH CDC20.
RX MEDLINE=98301442; PubMed=9637688;
RA Fang G., Yu H., Kirschner M.W.;
RT "The checkpoint protein MAD2 and the mitotic regulator CDC20 form a
RT ternary complex with the anaphase-promoting complex to control
RT anaphase initiation.";
RL Genes Dev. 12:1871-1883(1998).
RN [5];
RP INTERACTION WITH ADAM17.
RX PubMed=10527948;
RA Nelson K.K., Schlondorff J., Blobel C.P.;
RT "Evidence for an interaction of the metalloprotease-disintegrin tumour
RT necrosis factor alpha convertase (TACE) with mitotic arrest deficient
RT 2 (MAD2), and of the metalloprotease-disintegrin MDC9 with a novel
RT MAD2-related protein, MAD2-beta.";
RL Biochem. J. 343:673-680(1999).
RN [6];
RP STRUCTURE BY NMR OF 11-195.
RX MEDLINE=20165182; PubMed=10700282;
RA Luo X., Fang G., Coldiron M., Lin Y., Yu H., Kirschner M.W.,
```

```
RA Wagner G.;
RT "Structure of the Mad2 spindle assembly checkpoint protein and its
RT interaction with Cdc20.";
RL Nat. Struct. Biol. 7:224-229(2000).
CC -!- FUNCTION: REQUIRED FOR THE EXECUTION OF THE MITOTIC CHECKPOINT
CC WHICH MONITORS THE PROCESS OF KINETOCHORE-SPINDLE ATTACHMENT AND
CC DELAYS THE ONSET OF ANAPHASE WHEN THIS PROCESS IS NOT COMPLETE. IT
CC INHIBITS THE ACTIVITY OF THE ANAPHASE PROMOTING COMPLEX BY
CC SEQUESTERING CDC20 UNTIL ALL CHROMOSOMES ARE ALIGNED AT THE
CC METAPHASE PLATE.
CC -!- SUBUNIT: INTERACTS WITH CDC20 AND WITH ADAM17 (TACE).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE MAD2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U65410; AAC50781.1; -.
DR EMBL; U31278; AAC52060.1; -.
DR EMBL; AJ000186; CAA03943.1; -.
DR MIM; 601467; -.
DR PDB; 1DUJ; 08-MAR-00.
DR InterPro; IPR003511; -.
DR Pfam; PF02301; HORMA; 1.
KW Cell cycle; Mitosis; Nuclear protein; 3D-structure.
SQ SEQUENCE 205 AA; 23510 MW; B8DCBF0043836764 CRC64;

Query Match 52.6%; Score 40; DB 1; Length 205;
Best Local Similarity 54.5%; Pred. No. 6.2;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTF 11
   | | : | : | |
Db 33 YQGIYPSETF 43

RESULT 10
MD21_MOUSE
ID MD21_MOUSE STANDARD; PRT; 205 AA.
AC Q9Z1B5; Q9JI53;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MITOTIC SPINDLE ASSEMBLY CHECKPOINT PROTEIN MAD2A (MAD2-LIKE 1).
GN MAD2L1 OR MAD2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Jin D.-Y., Jeang K.-T.;
RT "Identification of a novel component of the spindle assembly
RT checkpoint in mammalian cells.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBBJ databases.
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX PubMed=10892650;
RA Dobles M., Liberal V., Scott M.L., Benezra R., Sorger P.K.;
RT "Chromosome missegregation and apoptosis in mice lacking the mitotic
RT checkpoint protein Mad2.";
RL Cell 101:635-645(2000).
CC -!- FUNCTION: REQUIRED FOR THE EXECUTION OF THE MITOTIC CHECKPOINT
CC WHICH MONITORS THE PROCESS OF KINETOCHORE-SPINDLE ATTACHMENT AND
CC DELAYS THE ONSET OF ANAPHASE WHEN THIS PROCESS IS NOT COMPLETE. IT
```



```
CC INHIBITS THE ACTIVITY OF THE ANAPHASE PROMOTING COMPLEX BY
CC SEQUESTERING CDC20 UNTIL ALL CHROMOSOMES ARE ALIGNED AT THE
CC METAPHASE PLATE (BY SIMILARITY).
CC -!- SUBUNIT: INTERACTS WITH CDC20.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MAD2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U83902; AAD09238.1; -.
DR EMBL; AF261919; AAF69525.1; -.
DR InterPro; IPR003511; -.
DR Pfam; PF02301; HORMA; 1.
KW Cell cycle; Mitosis; Nuclear protein.
FT CONFLICT 157 157 T -> A (IN REF. 1).
FT CONFLICT 178 178 C -> S (IN REF. 1).
FT CONFLICT 201 201 T -> I (IN REF. 1).
FT SEQUENCE 205 AA; 23598 MW; A9F3F28BC4C9738E CRC64;
SQ

Query Match 52.6%; Score 40; DB 1; Length 205;
Best Local Similarity 54.5%; Pred. No. 6.2;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTF 11
| | :||:|
Db 33 YQGIYPSETF 43

RESULT 11
CYSL_LYCES STANDARD; PRT; 346 AA.
AC P20721;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE LOW-TEMPERATURE-INDUCED CYSTEINE PROTEINASE PRECURSOR (EC 3.4.22.-)
DE (FRAGMENT).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. VFNT CHERRY;
RA Schaffer M.A., Fischer R.L.;
RT "Analysis of mRNAs that accumulate in response to low temperature
RT identifies a thiolprotease in tomato.";
RL Plant Physiol. 87:431-436(1988).
CC -!- INDUCTION: IN RESPONSE TO LOW TEMPERATURE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC PAPAIN FAMILY OF THIOL PROTEASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21444; AAA66308.1; -.
DR PIR; JA0159; JA0159.
DR HSSP; P00785; 2ACT.
DR MEROPS; C01.029; -.
DR InterPro; IPR000118; -.

```

```
DR InterPro; IPR000169; -.
DR InterPro; IPR000668; -.
DR Pfam; PF00112; peptidase C1; 1.
DR Pfam; PF00396; granulin; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.
KW Hydrolase; Thiol protease; zymogen; Glycoprotein.
FT NON_TER 1 1
FT PROPEP <1 17 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 18 346 LOW-TEMPERATURE-INDUCED CYSTEINE
FT PROTEINASE.
FT ACT_SITE 42 42 BY SIMILARITY.
FT ACT_SITE 178 178 BY SIMILARITY.
FT ACT_SITE 198 198 BY SIMILARITY.
FT DISULFID 39 81 BY SIMILARITY.
FT DISULFID 73 114 BY SIMILARITY.
FT DISULFID 172 223 BY SIMILARITY.
FT CARBOHYD 215 215 N-LINKED (GLCNAC...) (POTENTIAL).
FT SEQUENCE 346 AA; 37429 MW; D42AC73944010928 CRC64;
SQ

Query Match 52.6%; Score 40; DB 1; Length 346;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 MRPVYPTKTFPN 13
: | | | | |
Db 227 IEPSYPVKGTGP 238

RESULT 12
RRPO_PMV
ID RRPO_PMV STANDARD; PRT; 1547 AA.
AC P20951;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RNA REPLICATION PROTEIN (176 KDA PROTEIN) (ORF 1) [CONTAINS: RNA-
DE DIRECTED RNA POLYMERASE (EC 2.7.7.48); PROBABLE HELICASE].
OS Papaya mosaic potexvirus (PMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=12181;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89381685; PubMed=2778435;
RA Sit T.L., Abouhaidar M.G., Holy S.;
RT "Nucleotide sequence of papaya mosaic virus RNA.";
RL J. Gen. Virol. 70:2325-2331(1989).
CC -!- FUNCTION: RNA-REPLICATION. THE CENTRAL PART OF THIS PROTEIN
CC POSSIBLY FUNCTIONS AS A ATP-BINDING HELICASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D13957; BAA03050.1; -.
DR PIR; JQ0096; JQ0096.
DR InterPro; IPR000606; -.
DR Pfam; PF01443; Viral_helicase1; 1.
KW ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;
KW Transferrase.
FT NP_BIND 822 829 ATP (POTENTIAL).
FT SEQUENCE 1547 AA; 176320 MW; 6CA4282C6A082622 CRC64;
SQ

Query Match 52.6%; Score 40; DB 1; Length 1547;
Best Local Similarity 87.5%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```



QY 6 YPTKTFPN 13  
|||||  
Db 409 YPTKTFDN 416

RESULT 13  
HMD2\_SOLTU STANDARD; PRT; 595 AA.  
AC Q41437;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE 3-HYDROXY-3-METHYLGLUTARYL-COENZYME A REDUCTASE 2 (EC 1.1.1.34) (HMG-  
COA REDUCTASE 2) (HMG2.2).  
GN HMG2.  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
OC Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. KENNEBEC; TISSUE=Tuber;  
RX MEDLINE=97201488; PubMed=9049274;  
RA Korth K.L., Stermer B.A., Bhattacharyya M.K., Dixon R.A.;  
RT "HMG-CoA reductase gene families that differentially accumulate  
transcripts in potato tubers are developmentally expressed in floral  
tissues.";  
RL Plant Mol. Biol. 33:545-551(1997).  
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF MEVALONATE. THE SPECIFIC  
PRECURSOR OF ALL ISOPRENOID COMPOUNDS PRESENT IN PLANTS.  
CC -!- CATALYTIC ACTIVITY: (R)-MEVALONATE + COA + 2 NADP(+) = (S)-3-  
HYDROXY-3-METHYLGLUTARYL-COA + 2 NADPH.  
CC -!- PATHWAY: MEVALONATE BIOSYNTHESIS, PLANT ISOPRENOID BIOSYNTHESIS.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
RETICULUM.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN YOUNG FLOWERS AND IN MATURE  
SEEDS AND OVARIES.  
CC -!- SIMILARITY: BELONGS TO THE HMG-COA REDUCTASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U51985; AAB52551.1; -  
DR Mendel; 10574; Soltu; 1091; 10574.  
DR InterPro; IPR002202; -  
DR Pfam; PF00368; HMG-CoA\_red; 1.  
DR PRINTS; PR00071; HMGCOORDTASE.  
DR PROSITE; PS00066; HMG\_COA\_REDUCTASE\_1; 1.  
DR PROSITE; PS00318; HMG\_COA\_REDUCTASE\_2; 1.  
DR PROSITE; PS01192; HMG\_COA\_REDUCTASE\_3; 1.  
DR PROSITE; PS50065; HMG\_COA\_REDUCTASE\_4; 1.  
KW Oxidoreductase; Glycoprotein; Endoplasmic reticulum; Transmembrane;  
KW Isoprene biosynthesis; NADP; Multigene family.  
FT DOMAIN 1 112 MEMBRANE-BOUND (BY SIMILARITY).  
FT DOMAIN 113 183 LINKER (BY SIMILARITY).  
FT DOMAIN 184 595 CATALYTIC (BY SIMILARITY).  
FT TRANSMEM 48 68 POTENTIAL.  
FT TRANSMEM 92 112 POTENTIAL.  
FT ACT\_SITE 278 278 BY SIMILARITY.  
FT ACT\_SITE 486 486 BY SIMILARITY.  
FT ACT\_SITE 584 584 GENERAL BASE (BY SIMILARITY).  
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 342 342 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 588 588 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 595 AA; 63841 MW; 0FA7069849D41D57 CRC64;  
Query Match 51.3%; Score 39; DB 1; Length 595;  
Best Local Similarity 66.7%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 3 RPVYPTKTF 11  
:||||:|  
Db 9 KPVYPSKV 17

RESULT 14  
YCF2\_MESVI STANDARD; PRT; 890 AA.  
AC Q9MUP8;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL 103.9 KDA PROTEIN YCF2 (RF2).  
GN YCF2.  
OS Mesostigma viride.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;  
OC Chlorodendrales; Mesostigmataceae; Mesostigma.  
OX NCBI\_TaxID=41882;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NIES-296;  
RX MEDLINE=20150907; PubMed=10688199;  
RA Lemieux C., Otis C., Turmel M.;  
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early  
branch of green plant evolution.";  
RL Nature 403:649-652(2000).  
CC -!- SIMILARITY: BELONGS THE YCF2 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF166114; AAF43852.1; -  
DR InterPro; IPR001081; -  
DR InterPro; IPR001939; -  
DR Pfam; PF00004; AAA; 1.  
DR PROSITE; PS00674; AAA; 1.  
KW Chloroplast; Hypothetical protein.  
SQ SEQUENCE 890 AA; 103935 MW; .43CAEEEE991AF2C4B CRC64;

Query Match 51.3%; Score 39; DB 1; Length 890;  
Best Local Similarity 70.0%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 RPVYPTKTF 12  
|||||  
Db 650 RPVYSVKLFP 659

RESULT 15  
KIN1\_YEAST STANDARD; PRT; 1064 AA.  
AC P13185; Q04606;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROTEIN KINASE KIN1 (EC 2.7.1.-).  
GN KIN1 OR YDR122W OR YD9727.17.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

```
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87317589; PubMed=2957690;
RA Levin D.E., Hammond C.I., Ralston R.O., Bishop J.M.;
RT "Two yeast genes that encode unusual protein kinases.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:6035-6039(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Murphy L., Shore L., Harris D., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN IS PROBABLY A SERINE/THREONINE PROTEIN
CC KINASE.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIM1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M69017; AAA34722.1; -.
DR EMBL; Z48758; CAA88675.1; -.
DR PIR; S42438; S42438.
DR HSSP; Q63450; 1A06.
DR SGD; S0002529; KIN1.
DR InterPro; IPR000719; -.
DR InterPro; IPR002290; -.
DR Pfam; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 120..398 PROTEIN KINASE.
FT NP_BIND 126..134 ATP (BY SIMILARITY).
FT BINDING 149..149 ATP (BY SIMILARITY).
FT ACT_SITE 269..269 BY SIMILARITY.
FT CONFLICT 25..25 S -> R (IN REF. 1).
FT CONFLICT 453..453 T -> H (IN REF. 1).
FT CONFLICT 455..455 V -> G (IN REF. 1).
FT CONFLICT 718..718 A -> R (IN REF. 1).
FT CONFLICT 920..921 NI -> IN (IN REF. 1).
FT CONFLICT 976..976 T -> A (IN REF. 1).
FT CONFLICT 979..980 SI -> T (IN REF. 1).
FT CONFLICT 984..985 KT -> NS (IN REF. 1).
SQ SEQUENCE 1064 AA; 120070 MW; E4104B84A8E45C36 CRC64;
```

Query Match 51.3%; Score 39; DB 1; Length 1064;  
Best Local Similarity 50.0%; Pred. No. 52;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFP 12  
||||| | : |  
Db 802 YMRPPMPSSAYP 813

Search completed: July 19, 2001, 14:47:09  
Job time: 171 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2001, 14:46:48 ; Search time 42.51 Seconds  
(without alignments)  
40.460 Million cell updates/sec

Title: US-09-483-831-69\_COPY\_201\_213  
Perfect score: 76  
Sequence: 1 YMRPVYPTKTFPN 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues 425026  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_16:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_unclassified:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	862	11 Q9R1E6	Q9r1e6 mus musculu
2	76	100.0	863	4 Q13822	Q13822 homo sapien
3	76	100.0	863	4 Q15117	Q15117 homo sapien
4	76	100.0	915	4 Q13827	Q13827 homo sapien
5	70	92.1	885	11 Q64610	Q64610 rattus norv
6	69	90.8	257	11 Q88827	Q88827 rattus norv
7	69	90.8	845	4 Q9NP23	Q9np23 homo sapien
8	69	90.8	925	4 Q9P1P6	Q9p1p6 homo sapien
9	68	89.5	875	11 P97675	P97675 rattus norv
10	68	89.5	876	11 P70641	P70641 rattus norv
11	65	85.5	875	4 Q14638	Q14638 homo sapien
12	65	85.5	875	11 Q63490	Q63490 rattus norv
13	65	85.5	875	11 P97676	P97676 rattus norv
14	64	84.2	32	6 Q9TRD2	Q9trd2 bos taurus
15	64	84.2	817	14 Q90761	Q90761 fowlpox vir
16	64	84.2	817	14 Q9J5H1	Q9j5h1 fowlpox vir
17	49	64.5	457	10 Q9SU82	Q9su82 arabidopsis
18	49	64.5	461	10 Q9SU81	Q9su81 arabidopsis
19	49	64.5	829	5 P90754	P90754 caenorhabdi

20	48	63.2	479	10 Q9FS13	Q9fs13 spinacia ol
21	48	63.2	496	10 Q9SU83	Q9su83 arabidopsis
22	47	61.8	479	10 Q42974	Q42974 oryza sativ
23	45	59.2	674	5 P90755	P90755 caenorhabdi
24	43	56.6	173	4 Q9Y5A0	Q9y5a0 homo sapien
25	43	56.6	343	3 Q9UUN9	Q9uun9 sporobolomy
26	43	56.6	401	4 Q9HAC9	Q9hac9 homo sapien
27	43	56.6	757	11 Q9WVC8	Q9wvc8 mus musculu
28	42	55.3	429	2 Q69013	Q69013 zymomonas m
29	42	55.3	1004	5 Q9V3H9	Q9v3h9 drosophila
30	41	53.9	247	5 Q44145	Q44145 caenorhabdi
31	41	53.9	381	5 Q9XUV4	Q9xuv4 caenorhabdi
32	41	53.9	427	10 Q43423	Q43423 dianthus ca
33	41	53.9	433	2 Q9PAB9	Q9pab9 xylella fas
34	41	53.9	453	4 Q9Y6X5	Q9y6x5 homo sapien
35	41	53.9	493	10 Q22499	Q22499 zea mays (m
36	41	53.9	1829	5 Q22248	Q22248 caenorhabdi
37	40	52.6	203	13 P79935	P79935 xenopus lae
38	40	52.6	247	10 Q9SN23	Q9sn23 arabidopsis
39	40	52.6	291	10 Q9M4F6	Q9m4f6 brassica na
40	40	52.6	307	5 Q61478	Q61478 dreissena p
41	40	52.6	430	5 Q9NB54	Q9nb54 dreissena p
42	40	52.6	466	10 Q49877	Q49877 lycopersico
43	40	52.6	2066	14 P89920	P89920 oat blue dw
44	40	52.6	3579	5 Q9V5N8	Q9v5n8 drosophila
45	39	51.3	157	5 Q9Y0E8	Q9y0e8 drosophila

ALIGNMENTS

RESULT 1  
Q9R1E6  
ID Q9R1E6 PRELIMINARY; PRT; 862 AA.  
AC Q9R1E6;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE PHOSPHODIESTERASE I/NUCLEOTIDE PYROPHOSPHATASE (EC 3.1.4.1).  
GN NPFS2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57B/6;  
RA Sano K., Piao J.-H.;  
RT "Cloning and chromosomal assignment of mouse phosphodiesterase  
RT I/nucleotide pyrophosphatase (PD-I alpha/autotaxin).";  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF123542; AAD46480.1; -.  
DR InterPro; IPR001212; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR Pfam; PF01663; Phosphodiester; 1.  
DR PRINTS; PR00022; SOMATOMEDINB.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
DR SMART; SM00201; SO; 1.  
KW Hydrolase.  
SQ SEQUENCE 862 AA; 98918 MW; 641F8F1443E4F8B0 CRC64;

Query Match 100.0%; Score 76; DB 11; Length 862;  
Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
| | | | | | | | | | | | | | | | | |  
Db 200 YMRPVYPTKTFPN 212

RESULT 2





```
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX MEDLINE=95050605; PubMed=7961762;
RA Narita M., Goji J., Nakamura H., Sano K.;
RT "Molecular cloning, expression, and localization of a brain-specific
RT phosphodiesterase I/nucleotide pyrophosphatase (PD-I alpha) from rat
RL J. Biol. Chem. 269:28235-28242(1994).
DR EMBL; D28560; BAA05910.1; -.
DR InterPro; IPR001212; -.
DR InterPro; IPR001604; -.
DR InterPro; IPR002591; -.
DR Pfam; PF01033; Somatomedin_B; 2.
DR Pfam; PF01663; Phosphodiester; 1.
DR PRINTS; PR00022; SOMATOMEDINB.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
DR SMART; SM00477; NUC; 1.
KW Hydrolase.
SQ SEQUENCE 885 AA; 101309 MW; 8C5C0AFC52007973 CRC64;

Query Match          92.1%; Score 70; DB 11; Length 885;
Best Local Similarity 92.3%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13
Db 198 YTRPVYPTKTFPN 210

RESULT 6
O88827 ID O88827 PRELIMINARY; PRT; 257 AA.
AC O88827;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PLASMA CELL MEMBRANE GLYCOPROTEIN (FRAGMENT).
GN PC1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Sakoda H., Asano T., Ogiwara T.;
RT "N terminal of the plasma cell membrane glycoprotein PC-1.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017596; BAA33393.1; -.
DR InterPro; IPR001212; -.
DR Pfam; PF01033; Somatomedin_B; 2.
DR PRINTS; PR00022; SOMATOMEDINB.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
DR SMART; SM00201; SO; 1.
FT NON_TER 257 257
SQ SEQUENCE 257 AA; 28306 MW; CF348E474FBDAFED CRC64;

Query Match          90.8%; Score 69; DB 11; Length 257;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MRPVYPTKTFPN 13
Db 230 MRPVYPTKTFPN 241

RESULT 7
Q9NP23 ID Q9NP23 PRELIMINARY; PRT; 845 AA.
AC Q9NP23;
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE DJ131F15.2 (PHOSPHODIESTERASE I/NUCLEOTIDE PYROPHOSPHATASE 1
DE (HOMOLOGOUS TO MOUSE LY-41 ANTIGEN) (PC1, NPPS)) (FRAGMENT).
GN PDNP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Peck A.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL117378; CAB99365.1; -.
DR InterPro; IPR001212; -.
DR InterPro; IPR001604; -.
DR InterPro; IPR002591; -.
DR Pfam; PF01033; Somatomedin_B; 2.
DR Pfam; PF01663; Phosphodiester; 1.
DR PRINTS; PR00022; SOMATOMEDINB.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
DR SMART; SM00477; NUC; 1.
FT NON_TER 1 1
SQ SEQUENCE 845 AA; 96844 MW; C2DD58248BDAC52E CRC64;

Query Match          90.8%; Score 69; DB 4; Length 845;
Best Local Similarity 100.0%; Pred. No. 0.00092;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MRPVYPTKTFPN 13
Db 168 MRPVYPTKTFPN 179

RESULT 8
Q9P1P6 ID Q9P1P6 PRELIMINARY; PRT; 925 AA.
AC Q9P1P6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PDNP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bozzali M., Pizzuti A., Trischitta E.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF110304; AAF36094.1; -.
DR EMBL; AF110280; AAF36094.1; JOINED.
DR EMBL; AF110281; AAF36094.1; JOINED.
DR EMBL; AF110283; AAF36094.1; JOINED.
DR EMBL; AF110284; AAF36094.1; JOINED.
DR EMBL; AF110285; AAF36094.1; JOINED.
DR EMBL; AF110286; AAF36094.1; JOINED.
DR EMBL; AF110287; AAF36094.1; JOINED.
DR EMBL; AF110288; AAF36094.1; JOINED.
DR EMBL; AF110289; AAF36094.1; JOINED.
DR EMBL; AF110290; AAF36094.1; JOINED.
DR EMBL; AF110291; AAF36094.1; JOINED.
DR EMBL; AF110292; AAF36094.1; JOINED.
DR EMBL; AF110293; AAF36094.1; JOINED.
DR EMBL; AF110294; AAF36094.1; JOINED.
DR EMBL; AF110295; AAF36094.1; JOINED.
DR EMBL; AF110296; AAF36094.1; JOINED.
DR EMBL; AF110297; AAF36094.1; JOINED.
DR EMBL; AF110298; AAF36094.1; JOINED.
DR EMBL; AF110299; AAF36094.1; JOINED.
DR EMBL; AF110300; AAF36094.1; JOINED.
```

DR EMBL; AF110301; AAF36094.1; JOINED.  
DR EMBL; AF110302; AAF36094.1; JOINED.  
DR EMBL; AF110303; AAF36094.1; JOINED.  
DR InterPro; IPR001212; -.  
DR InterPro; IPR001604; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR Pfam; PF01663; Phosphodiesterase 1.  
DR PRINTS; PR00022; SOMATOMEDINB.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; 2.  
DR SMART; SM00477; NUC; 1.  
SQ SEQUENCE 925 AA; 104924 MW; 0ECAA063801CAFE8 CRC64;

Query Match 90.8%; Score 69; DB 4; Length 925;  
Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MRPVYPTKTFPN 13  
|||||  
Db 248 MRPVYPTKTFPN 259

RESULT 9  
P97675 PRELIMINARY; PRT; 875 AA.  
AC P97675;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE ALKALINE PHOSPHODIESTERASE.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RX MEDLINE=97250927; PubMed=9096610;  
RA Scott L.J., Delautier D., Meerson N.R., Trugnan G., Goding J.W.,  
RA Maurice M.;  
RT "Biochemical and molecular identification of distinct forms of  
RT alkaline phosphodiesterase I expressed on the apical and basolateral  
RT plasma membrane surfaces of rat hepatocytes.";  
RL Hepatology 25:995-1002(1997).  
DR EMBL; U78787; AAB61535.1; -.  
DR InterPro; IPR001212; -.  
DR InterPro; IPR001604; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR Pfam; PF01663; Phosphodiesterase 1.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; UNKNOWN\_2.  
DR SMART; SM00477; NUC; 1.  
SQ SEQUENCE 875 AA; 99166 MW; 8FEB9A482173D377 CRC64;

Query Match 89.5%; Score 68; DB 11; Length 875;  
Best Local Similarity 92.3%; Pred. No. 0.0014;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||||  
Db 197 YMRPVYPTKTFPN 209

RESULT 10  
P70641 PRELIMINARY; PRT; 876 AA.  
AC P70641;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE PHOSPHODIESTERASE I.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SMALL INTESTINE, PROXIMAL;  
RA Sano K.;  
RT "Molecular cloning of phosphodiesterase I cDNA from rat small  
RT intestine.";  
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D30649; BAA06333.1; -.  
DR InterPro; IPR001212; -.  
DR InterPro; IPR001604; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01033; Somatomedin\_B; 2.  
DR Pfam; PF01663; Phosphodiesterase 1.  
DR PROSITE; PS00524; SOMATOMEDIN\_B; UNKNOWN\_2.  
DR SMART; SM00477; NUC; 1.  
SQ SEQUENCE 876 AA; 99328 MW; D2F772C34A0C437A CRC64;

Query Match 89.5%; Score 68; DB 11; Length 876;  
Best Local Similarity 92.3%; Pred. No. 0.0014;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13  
|||||  
Db 197 YMRPVYPTKTFPN 209

RESULT 11  
O14638 PRELIMINARY; PRT; 875 AA.  
ID O14638;  
AC O14638;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE PHOSPHODIESTERASE I/NUCLEOTIDE PYROPHOSPHATASE BETA (EC 3.1.4.1)  
DE (PHOSPHODIESTERASE I BETA).  
GN PDNP3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RX MEDLINE=98008933; PubMed=9344668;  
RA Piao J.-H., Goding J.W., Nakamura H., Sano K.;  
RT "Molecular cloning and chromosomal localization of PD-Ibeta (PDNP3), a  
RT new member of the human phosphodiesterase I genes.";  
RL Genomics 45:412-415(1997).  
RN [2]  
RP SEQUENCE OF 189-875 FROM N.A.  
RA Hou S., Wohldmann P., Le T.;  
RT "The sequence of Homo sapiens PAC clone DJ0988G15.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 189-875 FROM N.A.  
RA Waterston R.H.;  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 189-875 FROM N.A.  
RA Waterston R.H.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 189-875 FROM N.A.  
RA Waterston R.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF005632; AAC51813.1; -.  
DR EMBL; AC005587; AAD05192.1; -.  
DR InterPro; IPR001212; -.  
DR

```
DR InterPro; IPR001604; -.
DR InterPro; IPR002591; -.
DR Pfam; PF01033; Somatomedin_B; 2.
DR Pfam; PF01663; Phosphodiester; 1.
DR PROSITE; PS00524; SOMATOMEDIN_B; UNKNOWN_2.
DR SMART; SM00477; NUC; 1.
KW Hydrolase.
SQ SEQUENCE 875 AA; 100096 MW; 329CF41667497BC4 CRC64;

Query Match      85.5%; Score 65; DB 4; Length 875;
Best Local Similarity 84.6%; Pred. No. 0.0048;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13
   ||| :|||||
Db 196 YMRAMYPTKTFPN 208

RESULT 12
Q63490 PRELIMINARY; PRT; 875 AA.
AC Q63490;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RB13-6 ANTIGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAQUE-DAWLEY; TISSUE=BRAIN;
RX MEDLINE=95247775; PubMed=7730366;
RA Deissler H., Lottspeich F., Rajewsky M.F.;
RT "Affinity purification and cDNA cloning of rat neural differentiation
RT and tumor cell surface antigen gp130RB13-6 reveals relationship to
RT human and murine PC-1.";
RL J. Biol. Chem. 270:9849-9855(1995).
DR EMBL; Z47987; CAA88029.1; -.
DR InterPro; IPR001212; -.
DR InterPro; IPR001604; -.
DR InterPro; IPR002591; -.
DR Pfam; PF01033; Somatomedin_B; 2.
DR Pfam; PF01663; Phosphodiester; 1.
DR PROSITE; PS00524; SOMATOMEDIN_B; UNKNOWN_2.
DR SMART; SM00477; NUC; 1.
SQ SEQUENCE 875 AA; 99087 MW; 9BA9FBFFEF82DFBA6 CRC64;

Query Match      85.5%; Score 65; DB 11; Length 875;
Best Local Similarity 84.6%; Pred. No. 0.0048;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13
   ||| :|||||
Db 197 YMRAMYPTKTFPN 209

RESULT 13
P97676 PRELIMINARY; PRT; 875 AA.
ID P97676;
AC P97676;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ALKALINE PHOSPHODIESTERASE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=97250927; PubMed=9096610;
RA Scott L.J., Delautier D., Meerson N.R., Trugnan G., Goding J.W.,
RA Maurice M.;
RT "Biochemical and molecular identification of distinct forms of
RT alkaline phosphodiesterase I expressed on the apical and basolateral
RT plasma membrane surfaces of rat hepatocytes.";
RL Hepatology 25:995-1002(1997).
DR EMBL; U78788; AAB61536.1; -.
DR InterPro; IPR001212; -.
DR InterPro; IPR002591; -.
DR Pfam; PF01033; Somatomedin_B; 2.
DR Pfam; PF01663; Phosphodiester; 1.
DR PROSITE; PS00524; SOMATOMEDIN_B; UNKNOWN_2.
DR SMART; SM00201; SO; 1.
SQ SEQUENCE 875 AA; 99071 MW; 4205F263E8A933EA CRC64;

Query Match      85.5%; Score 65; DB 11; Length 875;
Best Local Similarity 84.6%; Pred. No. 0.0048;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YMRPVYPTKTFPN 13
   ||| :|||||
Db 197 YMRAMYPTKTFPN 209

RESULT 14
Q9TRD2 PRELIMINARY; PRT; 32 AA.
AC Q9TRD2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE 260 KDA MAJOR ACIDIC FIBROBLAST GROWTH FACTOR-STIMULATED
DE PHOSPHOPROTEIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=94086550; PubMed=7505270;
RA Oda Y., Kuo M.D., Huang S.S., Huang J.S.;
RT "The major acidic fibroblast growth factor (aFGF)-stimulated
RT phosphoprotein from bovine liver plasma membranes has aFGF-stimulated
RT kinase, autoadenylation, and alkaline nucleotide phosphodiesterase
RT activities.";
RL J. Biol. Chem. 268:27318-27326(1993).
SQ SEQUENCE 32 AA; 3699 MW; 5CCFFA9EB55E6927 CRC64;

Query Match      84.2%; Score 64; DB 6; Length 32;
Best Local Similarity 91.7%; Pred. No. 0.0002;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MRPVYPTKTFPN 13
   ||||| |||
Db 2 MRPVYPTKXFPN 13

RESULT 15
O90761 PRELIMINARY; PRT; 817 AA.
ID O90761
AC O90761;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ALKALINE PHOSPHODIESTERASE I (EC 3.1.4.1) (5'-EXONUCLEASE) (5'-
DE NUCLEOTIDE PHOSPHODIESTERASE).
GN PC1.
OS Fowlpox virus.
```

OC Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Avipoxvirus.  
OX NCBI\_TaxID=10261;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FP-9;  
RX MEDLINE=98325194; PubMed=9658122;  
RA Laidlaw S.M., Anwar M.A., Thomas W., Green P., Shaw K., Skinner M.A.;  
RT "Fowlpox virus encodes nonessential homologs of cellular alpha-SNAP,  
RT PC-1, and an orphan human homolog of a secreted nematode protein.";  
RL J. Virol. 72:6742-6751(1998).  
CC -1- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES  
CC SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED  
CC OLIGO-NUCLEOTIDES.  
DR EMBL: AJ006408; CAA07014.1; -.  
DR InterPro; IPR001604; -.  
DR InterPro; IPR002591; -.  
DR Pfam; PF01663; Phosphodiester; 1.  
DR SMART; SM00477; NUC; 1.  
KW Hydrolase.  
SQ SEQUENCE 817 AA; 94004 MW; CB68DA0508CC568E CRC64;

Query Match 84.2%; Score 64; DB 14; Length 817;  
Best Local Similarity 91.7%; Pred. No. 0.0066;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MRPVPTKTFPN 13  
| | | | | | | | | |  
Db 155 MRPVPTNTFPN 166

Search completed: July 19, 2001, 14:46:49  
Job time: 172 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2001, 14:41:32 ; Search time 34.65 Seconds  
(without alignments)  
1600.893 Million cell updates/sec

Title: US-09-483-831-69  
Perfect score: 5019  
Sequence: 1 MARRSFQSCQIISLFTFAV.....RSYPEILTLKTYLHYESEI 915

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_0601.\*  
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5019	100.0	915	17	AAR86596	A2058 autotaxin pr
2	5019	100.0	915	22	AAAY71987	Human melanoma aut
3	4681	93.3	863	22	AAAY71988	Human teratocarcin
4	4632	92.3	859	22	AAAY71991	Human autotaxin pr
5	4533	90.3	829	17	AAR86578	Autotaxin derived
6	4430.5	88.3	858	22	AAAY71997	Rat autotaxin vari
7	4430.5	88.3	858	22	AAAY71998	Rat autotaxin vari
8	4429.5	88.3	858	22	AAAY71998	Rat autotaxin vari
9	4426.5	88.2	858	22	AAAY71986	Rat autotaxin prot
10	4423.5	88.1	858	22	AAAY71996	Rat autotaxin vari
11	4422.5	88.1	858	22	AAAY71995	Rat autotaxin vari

12	4341	86.5	849	17	AAR86595	N-tera 2D1 autotax
13	4306	85.8	885	22	AAAY71989	Rat brain autotaxi
14	4279	85.3	979	17	AAR86580	Autotaxin derived
15	3842	76.5	788	17	AAR86579	Autotaxin derived
16	1955.5	39.0	873	21	AAAY83620	Human PC-1 polypep
17	1955.5	39.0	925	16	AAR79148	Human insulin rece
18	1955.5	39.0	925	20	AAAY39355	Insulin receptor t
19	1955.5	39.0	925	21	AAB00195	Breast cancer prot
20	1954.5	38.9	873	21	AAAY83621	Variant human PC-1
21	646.5	12.9	457	21	AAG31321	Arabidopsis thalia
22	609	12.1	461	21	AAG29474	Arabidopsis thalia
23	609	12.1	465	21	AAG29473	Arabidopsis thalia
24	599	11.9	400	21	AAG29475	Arabidopsis thalia
25	569	11.3	359	21	AAG31322	Arabidopsis thalia
26	512	10.2	453	20	AAAY17529	Human secreted pro
27	512	10.2	453	20	AAAY02378	Polypeptide identi
28	512	10.2	453	22	AAB90545	Human secreted pro
29	484	9.6	338	21	AAG31323	Arabidopsis thalia
30	467	9.3	440	21	AAB18910	A novel polypeptid
31	465.5	9.3	477	21	AAB24072	Human PRO1107 prot
32	465.5	9.3	477	21	AAAY66716	Membrane-bound pro
33	465.5	9.3	477	22	AAB65239	Human PRO1107 (UNQ
34	465	9.3	411	20	AAAY34445	Porphyromonas ging
35	465	9.3	428	20	AAAY34324	Porphyromonas ging
36	360	7.2	438	21	AAW90962	Human CSGP-2 prote
37	357	7.1	355	19	AAW75859	Human secretory pr
38	357	7.1	355	20	AAAY45165	Human secreted pro
39	241.5	4.8	133	21	AAG31319	Arabidopsis thalia
40	200	4.0	92	21	AAG31320	Arabidopsis thalia
41	179	3.6	109	19	AAW52839	Secreted protein e
42	141	2.8	39	21	AAB53540	Human colon cancer
43	134	2.7	111	13	AAR26050	MSF-K130. Synthet
44	134	2.7	182	12	AAR10872	Protein encoded by
45	134	2.7	452	16	AAR80041	Human megakaryocyt

ALIGNMENTS

RESULT 1  
AAR86596  
ID AAR86596 standard; Protein; 915 AA.  
XX  
AC AAR86596;  
XX  
DT 01-JUL-1996 (first entry)  
XX  
DE A2058 autotaxin protein.  
XX  
KW Autotaxin; ATX; cytokine; autocrine motility stimulating protein; AMF;  
KW melanoma cell; tumour; antibody; cancer diagnosis; therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO9532221-A2.  
XX  
PD 30-NOV-1995.  
XX  
PF 24-MAY-1995; 95WO-US066613.  
XX  
PR 28-NOV-1994; 94US-0346455.  
PR 25-MAY-1994; 94US-0249182.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Krutzh J, Liotta L, Murata J, Schiffmann E, Stracke M;  
XX  
DR WPI; 1996-020533/02.  
XX  
PT Autotaxin motility stimulating protein, and DNA encoding it - used  
PT in cancer diagnosis and therapy  
XX  
PS Claim 4; Page 91-94; 112pp; English.

bad data

CC AAR86559-R86596 represent autotaxin (ATX) and fragments of it. This
 CC sequence represents the full length protein sequence of the A2058
 CC melanoma cell line ATX protein. ATX is an autocrine motility
 CC stimulating protein which is present in cancer cells. ATX stimulates
 CC both random and directed migration of melanoma cells. The tumorous form
 CC of ATX is a secreted protein, while the transmembrane bound form is not
 CC present in tumour cells. The cdna encoding this sequence can be used in
 CC a vector, to transform cells. The recombinant cells can then be used to
 CC produce the peptide sequences. Antibodies specific for these sequences
 CC can be produced, and can be used in cancer diagnosis and therapy.
 CC Different sites of localisation of the protein are utilised for diagnosis
 CC and prognosis of the stages of tumour progression. The sequences can be
 CC used in treatment methods to advantageously block the activity of the
 CC secreted form of AXT, while having little effect on the membrane form of
 CC AXT.
 XX
 SQ Sequence 915 AA;
 Query Match 100.0%; Score 5019; DB 17; Length 915;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MARRSSFQSCQIISLFTFVAVGVSICLGFHTAHRKRAEGWEEGPPTVLSDSPWTNISGSCK 60
 Db |
 QY 1 marrssfqcqisilftfavgvsiclgftahrikraegweegpptvlsdspwtnisgsk 60
 QY 61 GRCFELQEAAGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECKDRCGEVRNEENACHC 120
 Db |
 QY 61 grcfelqeagppdcrdnlcksytschdfdelclktargweckdrcgevrneenachc 120
 QY 121 SEDCLARGDCCTNYQVCKGESHVWDDCEEIKAAECPAGFVRPPLIIFSVGDFRASymK 180
 Db |
 QY 121 sedclargdcctnyqvckgeshwddceeiikaecpagfvrppliifsvdgrasymk 180
 QY 181 KGSVMNPNIKLRSCGTHSPYMRPVYTKTFPNLYTLATGLYPESHGIVGNSMYDPVfDA 240
 Db |
 QY 181 kgskvmnpnieklrscgthspymrpvytktfpnlytlatglypeshgivgnsmydpvfa 240
 QY 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVVIPHERILTLRLWLTLPDHER 300
 Db |
 QY 241 tfhlrgrekfnhrwggqplwitatkgvkagtfvswvipherriltlrlwltlpdher 300
 QY 301 PSVYAFYSEQPDFSGHKYGFPGPESSYGSFPTAKRPKRKVAPKRRQRPVAPPKKRRR 360
 Db |
 QY 301 psvyafyseqpdfsghkygpfpgpeesygsfptakrpkrvapkrqrqvappkrrr 360
 QY 361 KIHRMDHYAAETRODKMTNPLREIDKIVGQLMDGLKQLKLRRCNVNIFVGDHGMEDVTCD 420
 Db |
 QY 361 kihrmdhyaaetrqdkmntplreidkivgqlmdglkqlkrrcnvnvifvgdhgmedvtcd 420
 QY 421 RTEFLSNYLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKPDQHFkPYLKQH 480
 Db |
 QY 421 rteflsnyltnvdditlvpgtglgrirskfsnnakypkaiianltckpdqhfkylkqh 480
 QY 481 LPKRLHYANNRRRIEDIHLLVERRRWHVARKPLDVYKKPSGKCFQGDHGFDNKVNsmQTVF 540
 Db |
 QY 481 lpkrlhyannrrriedihllverrrwhvarkpldvyykkpsgkcfqgdhgfankvnsmqtvf 540
 QY 541 VGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLRLTNTFRPTmPEEV 600
 Db |
 QY 541 vgygptfkytkvppfenielynmcdllglkpapnngthgslnhlrltntfrptmpeev 600
 QY 601 TRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEEHLLYGRPAVLYRT 660
 Db |
 QY 601 trpnypgimylqsdfdlgctcddkvepknkldelnkrlhtkgsteerhllgrpavlyrt 660
 QY 661 RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVVRVSPFSQNC 720
 Db |
 QY 661 rydilyhtdfesgyseiflmlwtsytskqaevsvpdhltscvrpdvrvspfsqnc 720
 QY 721 AYKNDKQMSYGFLFPPYLLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYPQFVLVKKYASER 780

Db 721 aykndkqmsygfllppylssspeakydaflvtnmvmpafkrvwnyfqrvlvkkyaser 780
 QY 781 NGVNVISGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGP 840
 Db |
 QY 781 ngvnvisgpifdydydglhdtedkikqyvegssipvpthyysitscldftgpadkcdgp 840
 QY 841 LSVSSFILPHRPDNEESCNSSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTsrSYPE 900
 Db |
 QY 841 lsvssfilphrpdhneescnsdesekwveelkmhtarvrdiehltsldffrktssrype 900
 QY 901 ILTLKTYLHTYSEI 915
 Db |
 QY 901 iltlktylhtyesei 915
 RESULT 2
 AAY71987
 ID AAY71987 standard; Protein; 915 AA.
 XX
 AC AAY71987;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE Human melanoma autotaxin.
 XX
 KW Human; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;
 KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;
 KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;
 KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;
 KW type I phosphodiesterase activity; ATP pyrophosphatase activity;
 KW ATPase; adenosine-5'-triphosphatase activity; melanoma.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Binding-site 127..129
 FT /label= RGD\_binding\_domain
 FT Active-site 201..213
 FT /note= "Autotaxin phosphodiesterase catalytic site".
 XX
 PN WO200068386-A1.
 XX
 PD 16-NOV-2000.
 XX
 PF 05-MAY-2000; 2000WO-US12402.
 XX
 PR 07-MAY-1999; 99US-0306979.
 XX
 PA (ZYMO ) ZYMOGENETICS INC.
 XX
 PI Kelly JD;
 XX
 DR WPI; 2001-007397/01.
 XX
 PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and
 PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -
 XX
 PS Claim 27; Page 105-108; 126pp; English.
 XX
 CC The present sequence is autotaxin isolated from human melanoma
 CC cells. Autotaxin is a glycoprotein cytokine which increases
 CC insulin signalling in adipose tissue by producing substrate for
 CC adenosine receptors, resulting in inhibition of lipolysis,
 CC decreased hepatic gluconeogenesis and serum glucose levels, and
 CC increased insulin sensitivity. It also inhibits differentiation
 CC of adipocytes. Autotaxin has type I phosphodiesterase, adenosine-5'-
 CC triphosphatase (ATPase) and ATP pyrophosphatase activities.
 CC Autotaxin and its analogues are used to stimulate glucose uptake
 CC by cells, particularly to reduce serum glucose levels for
 CC treatment of non-insulin dependent diabetes (NIDDM) in humans,
 CC or generally any condition associated with elevated serum levels
 CC of glucose, lipid or free fatty acid (e.g. obesity or dyslipidemia).

CC Transgenic animals that overexpress autotaxin are models for  
CC human metabolic diseases.

XX Sequence 915 AA;  
SQ Query Match 100.0%; Score 5019; DB 22; Length 915;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 915; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARRSSFQSCQIISLFTFAVGVSICLFTAHRIKRAEGWEEGPPTVLSDSPWTNIGSCK 60  
Db 1 marrssfqcqlisltfavgvsiclgftahrikraegweegpptvlsdspwtnigsck 60  
QY 61 GRCFELQEAQPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGGEVNEENACHC 120  
Db 61 grcfelqeaqppdcrdnlcksytsccchdfdelclktagwectkdrcgevrneenachc 120  
QY 121 SEDCLARGDCCTNYQVCKGESHWVDDCEEIKAAECPAGFVRPPLIIFSVDFGRASYMK 180  
Db 121 sedclargdcctnyqvckgeshwvddceeikaaecpagfvrpplliifsvdgrasymk 180  
QY 181 KGSKVMPIEKLRSCTHSPYMRPVYPTKTPPNLYTLATGLYPESHGIVGNSMYDPVFDA 240  
Db 181 kgskvmnieklrscgthspymrpvyptktppnlytlatglypeshgivgnsmydpvfa 240  
QY 241 TFHLRGREKFNHRWGGQPLWITATKQGVKAGTFFWSVIPHERRILTLRLWTLPDHER 300  
Db 241 tfhlrgrekfnhrwggqplwitatkqgvkagttffwsvipherriltlrlwtlpdher 300  
QY 301 PSVYAFYSEQPDFSGHKYGPFGPEESSYSGPFTPAKRPRKKVAPKRRQRPVAPPKKRRR 360  
Db 301 psvyafyseqpdfsghkygpfpgpeessygsptpakrprkkrvapkrrrgerpvappkrrr 360  
QY 361 KIHMDHYAAETRODKMTNPLREIDKIVGQLMDGLKQLRRCVNVIFVGDHGMEDVTCD 420  
Db 361 kihrmdhyaaetrqdkmntplreidkivgqlmdglqklrrcvnvifvgdhgmedvtcd 420  
QY 421 RTEFLSNYLTNVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFPLYLKQH 480  
Db 421 rteflsnlyltnvdditlvpgtglgrirskfnnakypkaiianltckkpdqhfplylkqh 480  
QY 481 LPKRLHYANNRRIEDIHLVVERRHWVARKPLDVYKKPSGKCFQGDHGFDNKVNMQTVF 540  
Db 481 lpkrlhyannrriedihllverrhwvarkpldvyykpsgkcfqgdhgfdnkvnsmtqvf 540  
QY 541 VGYGPTFKYKTKVPPFENIELYNVCMCDLGLKPAPNNGTHGSLNHLRLTNRPTMPEEV 600  
Db 541 vgygptfkytkvppfenielynvcmcdlglkpapnngthgslnhlrltnrtprtmtpeev 600  
QY 601 TRPNYPGIMYLSDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEEERHLLYGRPAVLYRT 660  
Db 601 trpnypgimylsdfdlgctcddkvepknkldelnkrlhtkgsteerhllgrpavlyrt 660  
QY 661 RYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDRVSPFSQNCL 720  
Db 661 rydilyhtdfesgyseiflmlwtsytskqaevsvpdhltscvrpdvrvspfsqncl 720  
QY 721 AYKNDKQMSYGFLFPPYLYSSSPEAKYDAFLVTNMVPMYPAFKRWNVYFORVLVKKYASER 780  
Db 721 aykndkqmsygflfppylssspeakydaflvtnmvpmpafkrwnvyforvlvkkkyaser 780  
QY 781 NGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGP 840  
Db 781 ngvnvisgpifdydydglhdtedkikqyvegssipvpthyysiiitscldftqpadcddgp 840  
QY 841 LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPE 900  
Db 841 lsvssfilphrpdneescnsdeskwveelmkmhtarvrddiehltsldffrktsrsype 900  
QY 901 ILTLKTYLHTYESEI 915  
Db 901 iltlktylhtyesei 915

RESULT 3  
AAY71988  
ID AAY71988 standard; Protein; 863 AA.  
XX  
AC AAY71988;  
DT 28-MAR-2001 (first entry)  
XX Human teratocarcinoma autotaxin.  
DE Human; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; teratocarcinoma.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Binding-site 127..129  
FT /label= RGD\_binding\_domain  
FT Active-site 201..213  
FT /note= "Autotaxin phosphodiesterase catalytic site"

WO200068386-A1.

16-NOV-2000.

05-MAY-2000; 2000WO-US12402.

07-MAY-1999; 99US-0306979.

(ZYMO ) ZYMOGENETICS INC.

Kelly JD;

WPI; 2001-007397/01.

New autotaxin proteins, useful e.g. for treating diabetes mellitus and obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
Claim 27; Page 108-110; 126pp; English.

The present sequence is autotaxin isolated from human teratocarcinoma cells. Autotaxin is a glycoprotein cytokine which increases insulin signalling in adipose tissue by producing substrate for adenosine receptors, resulting in inhibition of lipolysis, decreased hepatic gluconeogenesis and serum glucose levels, and increased insulin sensitivity. It also inhibits differentiation of adipocytes. Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase (ATPase) and ATP pyrophosphatase activities. Autotaxin and its analogues are used to stimulate glucose uptake by cells, particularly to reduce serum glucose levels for treatment of non-insulin dependent diabetes (NIDDM) in humans, or generally any condition associated with elevated serum levels of glucose, lipid or free fatty acid (e.g. obesity or dyslipidemia). Transgenic animals that overexpress autotaxin are models for human metabolic diseases.

Sequence 863 AA;

Query Match 93.3%; Score 4681; DB 22; Length 863;  
Best Local Similarity 93.8%; Pred. No. 0;  
Matches 858; Conservative 2; Mismatches 3; Indels 52; Gaps 1;

QY 1 MARRSSFQSCQIISLFTFAVGVSICLFTAHRIKRAEGWEEGPPTVLSDSPWTNIGSCK 60  
Db 1 marrssfqcqlisltfavgvniclgftahrikraegweegpptvlsdspwtnigsck 60







Db 187 klrscgthspymrpvyptktpfnlytlatglypeshgivgnsmypvdfatfhlrgrekf 246

QY 251 NHRWGGQPLWITATKQGVKAGTFFWSVIPHERRILTLRWLTLPDHERPSVYAFYSEQ 310  
|||||

Db 247 nhrwggqplwitatkqgvkagttffwsvipherriiltqilwltlpdherpsvayfyseq 306

QY 311 PDFSGHKYGFPGPEESSYSGSPFTPAPKRKRKVAPKRQERPAPPKKRRRKIHRMDHYAA 370  
|||||

Db 307 pdfsgkhkygpfge----- 320

QY 371 ETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVDHGMEDVTCDRTEFLSNYLT 430  
|||||

Db 321 -----mtnplreidkivgqlmdgklqklhrcvnvifvgdhgmedvtcdrtflesnlyt 374

QY 431 NVDDITLVPGTGLGRIRSKFSNNAKYDPAKAIANLTCKKPDQHFKPYLKQHLPKRLHYANN 490  
|||||

Db 375 nvdditlvpgtigrirskfsnnakypkaiianltckkpdqhfkykqhlpkrlhyann 434

QY 491 RRIEDIHLLVERRHHVARKPLDVYKKPSGKCFQGDHGFNDKNVSMQTVFVGYPYFKYK 550  
|||||

Db 435 rriedihllverrhvarkpldvkykpsgkcfqgdhgfndknvsmqtvfvgypgpfkyk 494

QY 551 TKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEYTRPNYPGIMY 610  
|||||

Db 495 tkvppfenielynmcdllglkpapnngthgslnhllrtntfrptmpeeetrpnypgimy 554

QY 611 LQSDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRTRYDILYHTDF 670  
|||||

Db 555 lqsdfdlgctcddkvepknkldelnkrlhtkgsteerhlllygrpavlyrtrydilyhtdf 614

QY 671 ESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPFSQNCLAYKNDKQMSY 730  
|||||

Db 615 esgyseiflmlwtsytvskqaevssvpdhltscvrpdvrsvpsfsgncldaykndkqmsy 674

QY 731 GLFPYPYLSSSPEAKYDAFLVTNMVPMYPAFKRWVNYFQRLVLVKYASERNGVNVISGPI 790  
|||||

Db 675 glfpylsssspeakydaflvtnmvpmpafkrwnyfqrvlvkkyaserngvnvisgpi 734

QY 791 FDYDYGDLHDTEDIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGFLSVSSFILPH 850  
|||||

Db 735 fdydydglhdtedkikyvegssipvpthyysitscldftqpadkcdgplsvssfllph 794

QY 851 RPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHT 910  
|||||

Db 795 rpdneescnsdesekwveelmkmhtarvrddiehltsldffrktsrsypeiltlktylht 854

QY 911 YESEI 915

Db 855 yesei 859

RESULT 5

AAR86578

ID AAR86578 standard; Protein; 829 AA.

XX AC AAR86578;

XX 28-JUN-1996 (first entry)

DE Autotaxin derived from melanoma cell line A2058.

XX Autotaxin; ATX; cytokine; autocrine motility stimulating protein; AMF;

KW melanoma cell; tumour; antibody; cancer diagnosis; therapy.

XX Homo sapiens.

XX WO9532221-A2.

PN 30-NOV-1995.

PD 24-MAY-1995; 95WO-US06613.

XX 28-NOV-1994; 94US-0346455.

PR 25-MAY-1994; 94US-0249182.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA Krutzsch J, Liotta L, Murata J, Schiffmann E, Stracke M;

PI WPI; 1996-020533/02.

XX N-PSDB; TO6612.

DR Autotaxin motility stimulating protein, and DNA encoding it - used

DR in cancer diagnosis and therapy

XX Claim 4; Page 57-60; 112pp; English.

PS AAR86559-R86596 represent autotaxin (ATX) and fragments of it. ATX is

XX an autocrine motility stimulating protein which is present in cancer

CC cells. ATX stimulates both random and directed migration of melanoma

CC cells. The tumorous form of ATX is a secreted protein, while the

CC transmembrane bound form is not present in tumour cells. The cDNA

CC encoding this sequence can be used in a vector, to transform cells. The

CC recombinant cells can then be used to produce the peptide sequences.

CC Antibodies specific for these sequences can be produced, and can be used

CC in cancer diagnosis and therapy. Different sites of localisation of the

CC protein are utilised for diagnosis and prognosis of the stages of tumour

CC progression. The sequences can be used in treatment methods to

CC advantageously block the activity of the secreted form of AXT, while

CC having little effect on the membrane form of AXT.

XX Sequence 829 AA;

SQ

Query Match 90.3%; Score 4533; DB 17; Length 829;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 828; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 87 CHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARGDCCTNYQVVCKGESHWVD 146  
|||||

Db 1 chdfdelclktargwectkdrcevrneenachcsedclargdcctnyqvvcckgeshwvd 60

QY 147 DCEEIKAACECPAGFVRPPLIIFSDGFRASYMKKGSKVMPNIEKLRSCGTHSPYMRPVY 206  
|||||

Db 61 dceeikaacecpagfvrppliifsvdgrasymkkgskvmpnieklrscgthspymrpvy 120

QY 207 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKFNHRWGGQPLWITATK 266  
|||||

Db 121 ptktfpnlytlatglypeshgivgnsmypvdfatfhlrgrekfnhrwggqplwitatk 180

QY 267 QGVKAGTFFWSVIPHERRILTLRWLTLPDHERPSVYAFYSEQPDFSGHKYGFPGPEES 326  
|||||

Db 181 qgvkagtffwsvipherriltilrwltlpdherpsvayfyseqpdfsgkhkygpfgees 240

QY 327 SYGSPFTPAKRPRKRKVAPKRQERPAPPKKRRRKIHRMDHYAAETRQDKMTNPLREIDK 386  
|||||

Db 241 sygsptpakrpkrkvapkrqerpappkrrrkihrmdhyaaetrqdkmtnplreidk 300

QY 387 IVGQLMDGLKQLKLRRCVNVIFVDHGMEDVTCDRTEFLSNYITNVDDITLVPGTGLGRIR 446  
|||||

Db 301 ivgqlmdgklqlrrcvnvifvgdhgmedvtcdrtflesnlytnvdditlvpgtigrir 360

QY 447 SKFSNNAKYDPKAIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRRIEDIHLLVERRHWV 506  
|||||

Db 361 skfsnnakypkaiianltckkpdqhfkykqhlpkrlhyannrrriedihllverrhv 420

QY 507 ARKPLDVYKKPSGKCFQGDHGFNDKNVSMQTVFVGYPYFKYKTKVPPFENIELYNVMC 566  
|||||

Db 421 arkpldvkykpsgkcfqgdhgfndknvsmqtvfvgypgpfkyktkvppfenielynmvc 480

QY 567 DLLGLKPAPNNGTHGSLNHLRLTNTFRPTMPEEYTRPNYPGIMYLSQSDFDLGCTCDDKVE 626  
|||||

Db 481 dllglkpapnngthgslnhllrtntfrptmpeeetrpnypgimylsqsdldlgctcddkve 540

QY 627 PKNKLDLNLKRLHTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSY 686  
|||||

Db 541 pknkldelnlrhtkgtsteerhlllygrpavlyrtrydillyhtdfesgyseiflmlllwtsty 600

QY .687 TVSKQAEVSSVPDHLTSCVRPDVRVSPFSQNCLAYKNDQMSYGFLFPPYLSSSSPEAKY 746

Db 601 tvskqaevssvpdhltsctvrpdvrsvpsfsqncclaykndqmsygfllppylsssspeaky 660

QY 747 DAFLVTNMVMPYPAFKRVWNYFQRLVLVKKYASERNGVNVISGPIDFYDYDGLHDTEDKIK 806

Db 661 daflvtnmvmpyapfkrvwnyfqrvlvkkyaserngvnvvisgpidfydydglhdtedkik 720

QY 807 QYVEGSSIPVPTHYYSITSCLDFTQPADKCDGPLSVSSFILPHRPPDNEESCNSSEDESK 866

Db 721 qyvegssipvpthyysitscldftqpadkcdgplsvssfilphrpdneescnsdesk 780

QY 867 WVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915

Db 781 wveelmkmtarvrdiehltslldffrktsrsypeiltlktylhtyesei 829

RESULT 6

AA71997

ID AA71997 standard; Protein; 858 AA.

XX

AC AA71997;

XX

DT 28-MAR-2001 (first entry)

XX

DE Rat autotaxin variant (S236T).

XX

KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;

KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;

KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;

KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;

KW type I phosphodiesterase activity; ATP pyrophosphatase activity;

KW ATPase; adenosine-5'-triphosphatase activity; mutant; mutein; variant.

XX

OS Rattus sp.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 236

FT /note= "Wild type Ser substituted by Thr"

XX

PN WO200068386-A1.

XX

PD 16-NOV-2000.

XX

PF 05-MAY-2000; 2000WO-US12402.

XX

PR 07-MAY-1999; 99US-0306979.

XX

PA (ZYMO ) ZYMOGENETICS INC.

XX

PI Kelly JD;

XX

DR WPI; 2001-007397/01.

XX

PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and

PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -

XX

PS Disclosure; Page -; 126pp; English.

XX

CC The present sequence is variant (S236T) of rat autotaxin protein.

CC Autotaxin is a glycoprotein cytokine which increases insulin

CC signalling in adipose tissue by producing substrate for adenosine

CC receptors, resulting in inhibition of lipolysis, decreased hepatic

CC gluconeogenesis and serum glucose levels, and increased insulin

CC sensitivity. It also inhibits differentiation of adipocytes.

CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase

CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its

CC analogues are used to stimulate glucose uptake by cells, particularly

CC to reduce serum glucose levels for treatment of non-insulin dependent

CC diabetes (NIDDM) in humans, or generally any condition associated

CC with elevated serum levels of glucose, lipid or free fatty acid

CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress

CC autotaxin are models for human metabolic diseases.

CC Note: The present sequence is not shown in the specification

CC but is derived from rat autotaxin protein sequence found in

CC page 101-104 of sequence listing (AA71986).

XX

SQ Sequence 858 AA;

Query Match 88.3%; Score 4430.5; DB 22; Length 858;

Best Local Similarity 88.7%; Pred. No. 0;

Matches 803; Conservative 30; Mismatches 19; Indels 53; Gaps 2;

QY 11 QIISLFTFAVGYSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGCKGRCFELQEAG 70

Db 7 rvislftfaisvniclgtasrikrae-wdegpptvlstdspwtntsgsckgrcfelqevg 65

QY 71 PDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCEVRNEENACHCEDCLARGDC 130

Db 66 pdcrcdnlcksysscchdfdelclktargwectkdrcevrneenachsedclsrqdc 125

QY 131 CTNYQVVCKGESHWVDDDCCEEIKAAECPAGVRPPLIIFSVDFRASVMKKGSKVMPNIE 190

Db 126 ctnyqvvckgeshwvdddcceeikvpecpagvrpplifsvdgfrasmkkgskvmpnie 185

QY 191 KLRCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFHLRGREKF 250

Db 186 klrcgthapymrpvyptktnlytlatglypeshgivgnsmydpvdfhlrgrekf 245

QY 251 NHRWGGQPLWITATKQGVKAGTFFWSVVIHERRILTLRLWLTLPDHERPSVYAFYSEQ 310

Db 246 nhrwggqplwitatkqvragtffwsvispherriltilqlwslpdrpsvyafyseq 305

QY 311 PDFSGHKYGPFGPEESSYSPFTPAKRPKRKVAPKRQRQRPVAPPKRRRKIHRMDHYAA 370

Db 306 pdfsghkygpgfpe----- 319

QY 371 ETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCRTEFLSNYLT 430

Db 320 -----mtnplreidktvgqlmdgklqlhrccvnvifvgdhgmedvtcrtfslsnylt 373

QY 431 NVDDITLVPGTLGRIRSKFSNNAKYDPAKIIANLTCKPDQHFKPYLKOHLPKRLHYANN 490

Db 374 nvdditlvpgtlgriraksinnskydpktilanltckpdqhfkpykmqhlpkrlhyann 433

QY 491 RRIEDIHLLVERRHWVARKPLDVYKKPSGKCFQGDHGFNDKVNMQTFVFGYGPFTFYK 550

Db 434 rriedihllvdrwrhvarkpldvyykpsgkcfqgdhgfndkvnsmqtvfvgypftfkyr 493

QY 551 TKVPPFENIELYNVCMDDLGLKPAPNNGTHGSLNHLRLRTNFRPTMPEEVRPNYPGIMY 610

Db 494 tkvppfenielynvmcdllglkpapnngthgslnhlrltrntfrptmpdevsrpnypgimy 553

QY 611 LQSDFDLGCTCDDKVEPKNKLDELNKRHLHTKGSTEEHLLYGRPAVLRYTRYDILYHTDF 670

Db 554 lqsefdlgtctcdkvepknkleelnkrhlhtkgsteerhlllygrpavlyrtrydilyhtdf 613

QY 671 ESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPFSQNCLAYKNDQMSY 730

Db 614 esgyseiflmlwtsytiskqaevssipehltnrcvrdvrsvpgfsgncclaykndqmsy 673

QY 731 GFLFPPYLSSSSPEAKYDAFLVTNMVMPYPAFKRVWNYFQRLVLVKKYASERNGVNVISGPI 790

Db 674 gflfppylsssspeakydaflvtnmvmpyapfkrvwnyfqrvlvkkyaserngvnvvisgpi 733

QY 791 FDYDYGDLHDTEDKIKQYVEGSSIPVPTHYYSITSCLDFTQPADKCDGPLSVSSFILPH 850

Db 734 fdynydgldrtedeikqyvegssipvpthyysitscldftqpadkcdgplsvssfilph 793

QY 851 RPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHT 910

Db 794 rpdndescnsdeskweelmkmtarvrdiehltgldfyrktsrsypeiltlktylht 853

QY 911 YESEI 915  
Db 854 yesei 858

RESULT 7  
AAY71999  
ID AAY71999 standard; Protein; 858 AA.  
XX  
AC AAY71999;  
DT 28-MAR-2001 (first entry)  
XX  
DE Rat autotaxin variant (S289T).  
XX  
KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; mutant; mutein; variant.  
XX  
OS Rattus sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 289 /note= "Wild type Ser substituted by Thr"  
FT  
XX  
PN WO200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
XX  
DR WPI; 2001-007397/01.  
XX  
PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity. stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Disclosure; Page -: 126pp; English.  
XX  
CC The present sequence is variant (S289T) of rat autotaxin protein.  
CC Autotaxin is a glycoprotein cytokine which increases insulin  
CC signalling in adipose tissue by producing substrate for adenosine  
CC receptors, resulting in inhibition of lipolysis, decreased hepatic  
CC gluconeogenesis and serum glucose levels, and increased insulin  
CC sensitivity. It also inhibits differentiation of adipocytes.  
CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase  
CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its  
CC analogues are used to stimulate glucose uptake by cells, particularly  
CC to reduce serum glucose levels for treatment of non-insulin dependent  
CC diabetes (NIDDM) in humans, or generally any condition associated  
CC with elevated serum levels of glucose, lipid or free fatty acid  
CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress  
CC autotaxin are models for human metabolic diseases.  
CC Note: The present sequence is not shown in the specification  
CC but is derived from rat autotaxin protein sequence found in  
CC page 101-104 of sequence listing (AAY71986).  
XX  
SQ Sequence 858 AA;

Query Match 88.3%; Score 4430.5; DB 22; Length 858;  
Best Local Similarity 88.7%; Pred. NO. 0;  
Matches 803; Conservative 30; Mismatches 19; Indels 53; Gaps 2;

QY 11 QIISLTFEAVGVSIICLGFHTAHRIKRAEGWEEGPTVLSDSPWTNISGCKGRCFELQEAG 70  
Db :|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
7 rvislftfaisvniclgtasrikrae-wdegptvlstdspwtntsgsckgrcfelqevg 65

QY 71 PDCRCNDLCKSYTSCCHDFDELCLKTARGWECTKDRGCEVRNEENACHCSEDCLAGDC 130  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
66 pdcrcdnlcksyssccchdfdelclktargwectkdrcevrneenachcsedclsrqdc 125

QY 131 CTNYQVVCKGESHWVDDCEEIKAAECPAGFVRPPLIFSVDFGFRASYMKKGSKVMPNIE 190  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
126 ctnyqvvcckgeshwvddceeikvpecpagfvrppliifsvdgrasyymkkgskvmpnie 185

QY 191 KLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVDFATFHLRGREKF 250  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
186 klrscgthapymrpvyptktfnlytlatgypeshgivgnsmydpvdfasfhlrgrekf 245

QY 251 NHRWGGQPLWITATKQGVKAGTFFWSVIPHERIILTLRWLTLPDHERPSVYAFYSEQ 310  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
246 nhrwggqplwitatkqgvragtffwsvipheriiltilqwtltpdnerpsvyafyse 305

QY 311 PDFSGHKYGPFGPESSYGSPTPAKRPRKRVAPKRQRQRPVAPPKRRRKIRMDHYAA 370  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
306 pdfsgghkygpgfpe----- 319

QY 371 ETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNYLT 430  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
320 -----mtnplreidktvgqlmdglkqlrhrcvrvifvgdhgmmedvtdrteflsnlyt 373

QY 431 NVDDITLVPGTILGRIRSKFSNNAKYDPAKAIANLTCKKPDQHFYLPKQLPKRLHYANN 490  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
374 nvdditlvpgtigriraksinnskydpkaiianltckkpdqhfypmkqhlprkrlhyann 433

QY 491 RRIEDIHLLVERRWVARKPLDVYKKPSGKCFQGDHGFNDKVNMSQTVFVGYPGTFKYK 550  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
434 rriedihllvdrwrharkpldvyykpsgkcfqgdhgfndkvnsmqtvfvgypgtfkyr 493

QY 551 TKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLTLNTFRPTMPEEVRPNYPGIMY 610  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
494 tkvppfenielynmcdllglkpapnngthgslnhlrltntfrptmpdevsrpnypgimy 553

QY 611 LQSDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTERHLLYGRPAVLYRTRYDILYHTDF 670  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
554 lqsefdlgtctcddkvepknkleelnkrlhtkgsteerhlllygrpavlyrtsydilyhtdf 613

QY 671 ESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPFSQNCCLAYKNDKQMSY 730  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
614 esgyseiflmlwtsytiskvaevssipehltncvrpdvrsvspgfsqncclaykndkqmsy 673

QY 731 GFLEPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRLVLKKYASERNGVNVSIGPI 790  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
674 gflfppylssspeakydafvtnmvpmypafkrvwayfqrvlvkkyaserngvnvisgpi 733

QY 791 FDYDYGDLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPH 850  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
734 fdynydgldrtdedkikqyvegssipvpthyysiiitscldftqpadkcdgplsvssfilph 793

QY 851 RPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRYPEILTKTYLHT 910  
Db |:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |  
794 rpdndescnsedeskwveelmkmhtarvrddiehltdldfyrktsrypeiltktylht 853

QY 911 YESEI 915  
Db |:||||: |:||||: |:||||: |:||||: |:||||: |:||||: |:||||: |:||||: |  
854 yesei 858

RESULT 8  
AAY71998  
ID AAY71998 standard; Protein; 858 AA.  
XX  
AC AAY71998;  
XX  
DT 28-MAR-2001 (first entry)







FT Domain 5..15  
FT /note= "Putative transmembrane domain"  
FT Binding-site 122..124  
FT /label= RGD\_binding\_domain  
FT Region 143..158  
FT /label= Epitope  
FT /note= "This region is specifically claimed in claim 10"  
FT Region 149..158  
FT /label= Epitope  
FT /note= "This region is specifically claimed in claim 10;  
FT This region is absent in rat brain autotaxin designated  
FT as PD-Ialpha sequence (AA71989)"  
FT Active-site 196..208  
FT /note= "Autotaxin phosphodiesterase catalytic site"  
FT Region 585..595  
FT /label= Epitope  
FT /note= "This region is specifically claimed in claim 10"  
XX  
PN WO200068386-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 05-MAY-2000; 2000WO-US12402.  
XX  
PR 07-MAY-1999; 99US-0306979.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Kelly JD;  
DR WPI; 2001-007397/01.  
DR N-PSDB; AAD02131.  
XX  
PT New autotaxin proteins, useful e.g. for treating diabetes mellitus and  
PT obesity, stimulate glucose uptake by cells and inhibit lipolysis -  
XX  
PS Claim 1; Page 101-104; 126pp; English.  
XX  
CC The present sequence is rat autotaxin protein. Autotaxin is a  
CC glycoprotein cytokine which increases insulin  
CC signalling in adipose tissue by producing substrate for adenosine  
CC receptors, resulting in inhibition of lipolysis, decreased hepatic  
CC gluconeogenesis and serum glucose levels, and increased insulin  
CC sensitivity. It also inhibits differentiation of adipocytes.  
CC Autotaxin has type I phosphodiesterase, adenosine-5'-triphosphatase  
CC (ATPase) and ATP pyrophosphatase activities. Autotaxin and its  
CC analogues are used to stimulate glucose uptake by cells, particularly  
CC to reduce serum glucose levels for treatment of non-insulin dependent  
CC diabetes (NIDDM) in humans, or generally any condition associated  
CC with elevated serum levels of glucose, lipid or free fatty acid  
CC (e.g. obesity or dyslipidemia). Transgenic animals that overexpress  
CC autotaxin are models for human metabolic diseases.  
XX  
SQ Sequence 858 AA;

Query Match 88.2%; Score 4426.5; DB 22; Length 858;  
Best Local Similarity 88.6%; Pred. No. 0;  
Matches 802; Conservative 31; Mismatches 19; Indels 53; Gaps 2;

QY 11 QIISLFTFAVGVSICLGFTHARIKRAEGWEEGPPFTVLSDSPWTNISGCKGRCFELQEAG 70  
:|||||: |:||||| |:||||| |:||||| |:||||| |:||||| |:||||| |:|||||  
Db 7 rvislftfaisvniclghtasrikrae-wdegpptvlstdspwtntsgckgrcfelgev 65

QY 71 PPDRCNDNLCKSYTSCCHDFDELCLKTARGWECTKDRGCEVRNEENACHCSEDCLARGDC 130  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 66 ppdrcndnlckysccchdfdelclktargwectkdrcevrneenachcsedclsrge 125

QY 131 CTNYQVVCKGESHWDDDCIEIKAAECAPAGFVRPPLIFSDGFRASYMKKGSKVMPNIE 190  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 126 ctnyqvvckgeshwdddcieikvpecpagfvrppllifsvdgrasymkkgskvmpnie 185

QY 191 KLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGREKF 250

Db 186 klrscgthapymrvpyptktnlytlatglypeshgivgnsmydpvdfasflrgrekf 245  
QY 251 NHRWGGQPLWTATKQGVKAGTFFWSVVIPHERILTILRWLTLPDHERPSVYAFYSEQ 310  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 246 nhrwggqplwtatkqgvragtffwsvsipherriltilqlwslpdpnerpsvyafyse 305  
QY 311 PDFSGHKYGPFGPESSYGSPTPAKRPKRKVAPKRRQERPAPPKRRRKTHRMDHYAA 370  
|||||:|||||  
Db 306 pdfsghkygpgpe----- 319  
QY 371 ETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNYLT 430  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 320 -----mtnplreidktvgqlmdgklqlrlhrcvnnvifvgdhgmedvtcdrtfslsnylt 373  
QY 431 NVDDITLVPGTGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHYANN 490  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 374 nvdditlvpgtligriraksinnskydpktilianltckkpdqhfkpymkqhlpkrlhyann 433  
QY 491 RRIEDIHLLVERRHWVARKPLDVYKKPSGKCFQGDHGFDNKVNMSQTVFVGYPGTFKYK 550  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 434 rriedihllvdrrwhvarkpldvyykkpsgkcfqgdhgfdnkvnsmtqtfvgygptfkyr 493  
QY 551 TKVPPFENIELYNMCDLLGLKPAPNNGTHGSLNHLRTNTFRPTMPEEVRPNYPGIMY 610  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 494 tkvppfenielynvmcddlglkpapnngthgslnhlrtntfrptmpdevsrpnypgimy 553  
QY 611 LQSDFDLGCTCDDKVEPKNKLDLKNKRLHTKGSTEERHLLYGRPAVLRYTRDYLYHTDF 670  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 554 lqsefdlgctcddkvepknkleelnkrlhtkgsteerhlllygrpavlyrtsyailyhtdf 613  
QY 671 ESGYSEIFLMLLWTSYTVSKQAEVSSVDPDLTSCVRPDVRVSPFSQNCCLAYKNDKQMSY 730  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 614 esgyseiflmlwtsytiskqaevsispehltnrcvrdvrvspgfsqncclaykndkqmsy 673  
QY 731 GFLFPPYLSSSPEAKYDAFLVTNMVPMYPAPKRWVNYFORVLVKKYASERNGVNVISGPI 790  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 674 gflfpylssspeakydafvtnmvpmypapkrvwayfgrvlvkkyaserngvnvisgpi 733  
QY 791 FDYDYGDLHDTEDKIKQYVEGSSIPVPTHYYSITSCLDFTQPADKCDGPLSVSSFILPH 850  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 734 fdynydgldrtdedekqyvegssipvpthyysitscldftqpadkcdgplsvssfilph 793  
QY 851 RPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFKTSRSYPEILTKTYLHT 910  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 794 rpdndescnsdesedskwveelkmhtarvrdiehltdgldfyrktsrsyseilktylht 853  
QY 911 YESEI 915  
|||||  
Db 854 yesei 858

RESULT 10  
AA71996  
ID AA71996 standard; Protein; 858 AA.  
XX  
AC AA71996;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE Rat autotaxin variant (A194V).  
XX  
KW Rat; autotaxin; lipolysis; gluconeogenesis; insulin sensitivity;  
KW adipocyte; hypoglycemic; antidiabetic; anorectic; antilipemic;  
KW therapy; metabolic disease; obesity; dyslipidemia; serum glucose;  
KW non-insulin dependent diabetes mellitus; NIDDM; transgenic animal;  
KW type I phosphodiesterase activity; ATP pyrophosphatase activity;  
KW ATPase; adenosine-5'-triphosphatase activity; mutant; mutein; variant.  
OS Rattus sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers











Db 1 marqgclgsfqvislftfaisvniclgtasrikrae-wdegpptvlsdpswntsgsck 59  
QY 61 GRCFELQEAQPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKRCGEVRNEENACHC 120  
Db 60 grcfelqevgppdcrcdnlcksysccchdfdelclktvrgwectkdrsgvrneenachc 119  
QY 121 SEDCLARGDCCTNYQVVCKGESHVDD-----DCEIKAACECPAGFVRPPLIIFSVDG 173  
Db 120 pedclsrqdcctnyqvvcckgeshwddaarnqsseclqv---cp-----pplifsvdg 170  
QY 174 FRASYMKKGSKVMPNIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSM 233  
Db 171 frasymkkgskvmpnieklrscgthvpytrpvpyptktfpnllytlatgлыпeshgivgns 230  
QY 234 YDPVFDATFHLRGREKFNHRWGGOPLWITATKQGVKAGTFFWSVVIPHERRILTLRWL 293  
Db 231 ydpvfda fhlrgrekfnhrwggqplwitatkqvragtffwsvipherri ltilqlw1 290  
QY 294 TLPDHERPSVYAFYSEQPDFSGHKYGFPGPEESSYSGSPFPAKRPKRKVAPKRRQERPVA 353  
Db 291 slpdnerpsvyafyseqpdfsgkhygpgfpe----- 321  
QY 354 PPKKRRRKIHRMDHYAAETRODKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHG 413  
Db 322 -----mtnplreidktvgqlmdglkqlrlhrcvnvifvgdhg 358  
QY 414 MEDVTCDRTEFLSNYLTNVDDITLVPGTGLRIRSKFSNNAKYDKRAIIANLTCKKPDQHF 473  
Db 359 medvtcdrtelfsnlyltnvdditlvpgtlgraksinnskypdktiianltckkpdqhf 418  
QY 474 KPYLKQHLPKRLHYANNRRIEDIHLVERRWHVARKPLDVYKKPSKCFQGDHGFNDKV 533  
Db 419 kpymkqhlpkrlhyannrriedihlvdrwhvarkpldvkykpsgkcfqgdhgfndkv 478  
QY 534 NSMQTVFVGYGPTFKYKTKVPPFENIELYNMCDLLGLKPAPNNGTHGSLNHLLRTNTER 593  
Db 479 nsmqtvfvgygptfk yrtkvppfenielynmvcdllglkpapnngthgslnhllrtntfr 538  
QY 594 PTMPEEVRPNYPGIMYLOSDFDLGCTCDDKVEPKNKLDELNKRHLTKGSTE----- 645  
Db 539 ptmpdevsrpnypgimylqsefdlgctcddkvepknklee lnkrlhtkgsteaetgkfrg 598  
QY 646 -----ERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTV 688  
Db 599 skhenknlngsvprkerhllgrpavlyrtsy dilyhtdfesgyseiflmp lwtstyti 658  
QY 689 SKQAEVSSVPDHLTSCVRPDVRSVPSFSONCLAYKNDKQMSYGLFPPYLSSSPEAKYDA 748  
Db 659 skqae vssipehltncvrpdvrsvpgfsqnc laykndkmsyglfppylssspeakyda 718  
QY 749 FLVTNMVPMYPAFKRVWNYFQVLVKKYASERNGVNVISGPIDFYDYGDLHDTEDKIKQY 808  
Db 719 flvtnmvmpyafkrvwayfgrvlvkkyaserngvnv isgpidfyndy dglrdtedeikqy 778  
QY 809 VEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWV 868  
Db 779 vegssipvpthyysiitscldftqpadkcdgplsvssfilphrpndescnsse deskwv 838  
QY 869 EELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYSEI 915  
Db 839 eelmkmhtarvrdiehl tgdfyrtksrsyseil tkltylhtysei 885

RESULT 14  
AAR86580

ID AAR86580 standard; Protein; 979 AA.

XX AAR86580;

AC AAR86580;

XX 28-JUN-1996 (first entry)

DE Autotaxin derived from human liver cells.

XX Autotaxin; ATX; cytokine; autocrine motility stimulating protein; AMF;  
KW melanoma cell; tumour; antibody; cancer diagnosis; therapy.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 864 /note= "unspecified amino acid"  
FT Misc-difference 889 /note= "unspecified amino acid"  
FT Misc-difference 905 /note= "unspecified amino acid"  
FT Misc-difference 911 /note= "unspecified amino acid"  
FT Misc-difference 927 /note= "unspecified amino acid"  
FT Misc-difference 937 /note= "unspecified amino acid"  
FT Misc-difference 944 /note= "unspecified amino acid"  
FT Misc-difference 950 /note= "unspecified amino acid"  
FT Misc-difference 954 /note= "unspecified amino acid"  
FT Misc-difference 967 /note= "unspecified amino acid"  
FT Misc-difference 975 /note= "unspecified amino acid"  
FT Misc-difference 975 /note= "unspecified amino acid"  
XX WO9532221-A2.  
XX 30-NOV-1995.  
XX 24-MAY-1995; 95WO-US06613.  
XX 28-NOV-1994; 94US-0346455.  
XX 25-MAY-1994; 94US-0249182.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Krutzch J, Liotta L, Murata J, Schiffmann E, Stracke M;  
XX WPI; 1996-020533/02.  
XX Autotaxin motility stimulating protein, and DNA encoding it - used  
XX in cancer diagnosis and therapy  
XX Claim 4; Page 67-70; 112pp; English.  
XX AAR86559-R86596 represent autotaxin (ATX) and fragments of it. ATX is  
XX an autocrine motility stimulating protein which is present in cancer  
XX cells. ATX stimulates both random and directed migration of melanoma  
XX cells. The tumorous form of ATX is a secreted protein, while the  
XX transmembrane bound form is not present in tumour cells. The cDNA  
XX encoding this sequence can be used in a vector, to transform cells. The  
XX recombinant cells can then be used to produce the peptide sequences.  
XX Antibodies specific for these sequences can be produced, and can be used  
XX in cancer diagnosis and therapy. Different sites of localisation of the  
XX protein are utilised for diagnosis and prognosis of the stages of tumour  
XX progression. The sequences can be used in treatment methods to  
XX advantageously block the activity of the secreted form of AXT, while  
XX having little effect on the membrane form of AXT.  
XX Sequence 979 AA;

Query Match 85.3%; Score 4279; DB 17; Length 979;

Best Local Similarity 87.1%; Pred. No. 0;

Matches 805; Conservative 13; Mismatches 36; Indels 70; Gaps 8;

QY 1 MARRSSFQSQIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGCK 60

||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

Db 1 marrssfscqdislftfavgvniclgtahrikraegweegpptvlstdspwtntissgsk 60

QY 61 GRCFELQEAQPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRGGEVRNEENACHC 120

Db 61 grcfelqeagppdcrcdnlcksytsccchdfdelclktarawectkdrcevrneenachc 120

QY 121 SEDCLARGDCCTNYQVCKGESHVWDDCEEIKAAECFAGFVRPPLIIFSVDFG--RASY 178

Db 121 sedclargdcctnyqvckgeshwddcceikaaecfq--vcspsinhlrlrgwlpmtsy 178

QY 179 MKKGSKVMPIEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVF 238

Db 179 mkkgskvmpnieklrscgthspymrpvyptktfpnllylatglypeshgivgnsmydpvf 238

QY 239 DATFHLRGREKENHRWGGQPLWITATKQGVKAGTFEWSVVI----PHERRILFILRWLT 294

Db 239 datfhlrgrekfhrwggqplwitatkrges----wnillvcchpsraeiltlqlwt 294

QY 295 LPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSFPTPAKRPRKRVAPKRRRQERVPAP 354

Db 295 lpdherpsvayfseqpdfsgkhmpfge----- 324

QY 355 PKKRRRKIHRMDHYAAETRDQKMTNPLREIDKIVGQLMDGLKQLKRRRCVNVIFVGDHGM 414

Db 325 -----mtnplremhkivgqlmdgkqlklhrcvnvifv---et 359

QY 415 EDVTCD--RTEFLSNLYTNVDDITLVPGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQH 472

Db 360 mdgrchmyrteflsltnvdditlvpgtlgrirskfsnnakypkaiianltckkpdqh 419

QY 473 FKPYLKQHLPKRLHYANNRRIEDIHLLVERRHWVARKPLDVYKKPSGCKFFQGDHGFENK 532

Db 420 fkpylkqhlpkrlhyannrriedihllverrhvarkpldvkykpsgnafsrattadnk 479

QY 533 VNSMQTVFVGYPGTFKYTKVPPFENIELYNVMDLLGLKPAPNNGTHGSLNHLRTNTE 592

Db 480 vnsmtqvfvgypgtfkytkvppfenielynmcdllgkpaapnngthgslnhlrrtnf 539

QY 593 RPTMPEVTRPNYPGIMYLSQDFDLGCTCDDKVEPKNKLDLNLKRLHTKGSTEERHLLYG 652

Db 540 rptmpeevtrpnypgimylsqdfdlgctcddkvepknkldelnkrlhtkgsteerhllvg 599

QY 653 -RPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRV 711

Db 600 drpavlyrtrydilyhtdfesgyseiflmlplwtsytvskqaevsvpdhltscvrpdvr 659

QY 712 SPFSQONCLAYKNDQMSYGFLEPPYLSSSSPEAKYDAFLVTNMVMPYPAFKRVWNYFQV 771

Db 660 spfsqncalaykndqmsygfleppylssspeakydaflvtnmvmypafkrvwnyfqrv 719

QY 772 LVKKYASERNGVNVTSGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSLDFT 831

Db 720 lvkkyaserngvnvtsgpifdydglhdtedkikgyvegssipvpthyysitclft 779

QY 832 QPADKCDGPLSVSSFILPHRPDNEESCNSSEDESQWVEELMKMHTARVRDIEHLTSLDFF 891

Db 780 qpadkcdgplsvssfilphrpdneescnsdeskweelmkmtarvrddiehltsldff 839

QY 892 RKTSRYPEILTLKTYLHTYSEI 915

Db 840 rktsrypeiltlktylhtyesei 863

RESULT 15

AAR86579

ID AAR86579 standard; Protein; 788 AA.

XX

AC AAR86579;

XX

XX 28-JUN-1996 (first entry)

DT

XX Autotaxin derived from teratocarcinoma N-tera 2D1 cells.

DE

XX

KW Autotaxin; ATX; cytokine; autocrine motility stimulating protein; AMF; melanoma cell; tumour; antibody; cancer diagnosis; therapy.

KW

OS Homo sapiens.

XX

PN WO9532221-A2.

XX

PD 30-NOV-1995.

XX

PF 24-MAY-1995; 95WO-US066613.

XX

PR 28-NOV-1994; 94US-0346455.

PR 25-MAY-1994; 94US-0249182.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Krutzsch J, Liotta L, Murata J, Schiffmann E, Stracke M;

XX

DR WPI; 1996-020533/02.

DR N-PSDB; AAT06613.

XX

PT Autotaxin motility stimulating protein; and DNA encoding it - used in cancer diagnosis and therapy

XX

PS Claim 4; Page 62-65; 112pp; English.

XX

CC AAR86559-R86596 represent autotaxin (ATX) and fragments of it. ATX is an autocrine motility stimulating protein which is present in cancer cells. ATX stimulates both random and directed migration of melanoma cells. The tumorous form of ATX is a secreted protein, while the transmembrane bound form is not present in tumour cells. The cDNA encoding this sequence can be used in a vector, to transform cells. The recombinant cells can then be used to produce the peptide sequences. Antibodies specific for these sequences can be produced, and can be used in cancer diagnosis and therapy. Different sites of localisation of the protein are utilised for diagnosis and prognosis of the stages of tumour progression. The sequences can be used in treatment methods to advantageously block the activity of the secreted form of AXT, while having little effect on the membrane form of AXT.

CC

SQ Sequence 788 AA;

Query Match 76.5%; Score 3842; DB 17; Length 788;

Best Local Similarity 85.6%; Pred. No. 0;

Matches 727; Conservative 12; Mismatches 40; Indels 70; Gaps 8;

QY 76 CDNLCKSYTSCCHDFDELCLKTARGWECTKDRGGEVRNEENACHCEDCLARGDCCTNYQ 135

Db 1 cdnlcksytsccchdfdelclktarawectkdrcevrneenachcedclargdcctnyq 60

QY 136 VVCKGESHWVDDCEEIKAAECFAGFVRPPLIIFSVDFG--RASVMKKGSKVMPIEKLR 193

Db 61 vvckgeshwvddcceikaaecfq--vdspsinhllrgwlpmtsymkkgskvmpnieklr 118

QY 194 SCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFATFHLRGREKFNHR 253

Db 119 scgthspymrpvyptktfpnllylatglypeshgivgnsmydpvfatfhlrgrekfhn 178

QY 254 WGGQPLWITATKQGVKAGTFEWSVVI----PHERRILTLRWLTLPDHERPSVYAFYSE 309

Db 179 wwaggplwitatkrges----wnillvcchpsraeiltlqlwtlpdherpsvayfse 234

QY 310 QPDFSGHKYGPFGPEESSYGSFPTPAKRPRKRVAPKRRRQERVPAPKRRRKIHRMDHYA 369

Db 235 qpdfsghkmpfge----- 249

QY 370 AETRODKMTNPLREIDKIVGQLMDGLKQLKRRRCVNVIFVGDHGMEDVTCD--RTEFLSN 427

Db 250 -----mpnplremhkivgqlmdgkqlklhrcvnvifv---etmdgrchmyrtefln 299

QY 428 YLTNVDDITLVPGLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFEPYLKQHLPKRLHY 487

Db 300 yltnvdditlvpgtlgrirskfsnnakypkaiianltckkpdqhfkykqhlpkrlhy 359

Qy 488 ANRRRIEDHLLVERRWHVARKPLDVYKKPSGKCFQGDHGFNDKNVSMQTVFVGPTF 547  
|||||

Db 360 annrriedhliverrwvharkpldvkykpsgnafsrrettafdnkvnsmtvfvgygptf 419  
|||||

Qy 548 KYTKVPPFENIELYNVMDLLGLKPPAPNNGTHGSINHLLRTNFRPTMPEEVRPNYPG 607  
|||||

Db 420 kytkvppfenielynvmcdllglkppapnngthfslnhllrtntfrptmpeevtrpnypg 479  
|||||

Qy 608 IMYLSDFDLGCTCDDKVEPKNKLDLKNRHLHTKSGSTEERHLLYG-RPAVLYRTRYDILY 666  
|||||

Db 480 imylqsdfdlgctcdckvepknkldelnkrlhtkgsteerhlllygdrpavlyrtrydily 539  
|||||

Qy 667 HTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPSFSQNCLAYKNDK 726  
|||||

Db 540 htdfesgyseiflmpwtstvtvskqaevsvpdhltscvrpdvrvspfsqncclaykndk 599  
|||||

Qy 727 QMSYGFLFPYLSSSPEAKYDAFLVTNMVMPYPAKRVWNYFQRLVKKYASERNGVNVI 786  
|||||

Db 600 qmsygglgppylssspeakydaflvtnmvpmpafkrwnyfgvrlvkkyaserngvnvi 659  
|||||

Qy 787 SGPIFDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSF 846  
|||||

Db 660 sgpfidydyglhdtedkikqyvegssipvpthyysitscldftqpadkcdgplsvssf 719  
|||||

Qy 847 ILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRYPEILTLKT 906  
|||||

Db 720 ilphrpdneescnsdesekwveelmkmhtarvrddiehltsldffrktsrypeiltlkt 779  
|||||

Qy 907 YLHTYESEI 915  
|||||

Db 780 ylhtyesei 788